

筑波大学大学院博士課程
理工情報生命学術院
システム情報工学研究群修士論文

オンラインホワイトボードを用いた
遠隔会議における非言語情報の自動検出
および絵文字を用いた自動提示

大山 智弘
修士（工学）
（情報理工学位プログラム）

指導教員 志築 文太郎

2024年3月

概要

オンライン会議では一般的にビデオによる顔映像の共有を用いることにより非言語情報の補完が行われている。しかし、オンラインホワイトボードの共同編集を行いながらのオンライン会議では、ビデオを見ることが困難となる場合がある。そのような問題を解決するため本研究では、オンラインホワイトボードを用いて行われるオンライン創造会議にて、ユーザの顔映像からリアクションを自動取得し、絵文字化してカーソル位置とともに共有する手法を提案する。提案手法では、絵文字を用いることによりビデオよりも表示領域が小さく、既存のリアクションボタンを用いる手法のように意図的な操作を必要とせず非言語情報の伝達が可能になる。提案手法の効果を明らかにするために、2回の実験を行った。実験1では、現在広く用いられている手法と比較するため、提案手法とリアクションボタンを用いる既存手法、右側にビデオを表示する条件の3条件で比較実験を行った。結果として、提案手法では既存手法よりも多くのリアクションが共有されたこと、会議に対する満足度は既存手法よりも提案手法の方が高く、提案手法とビデオの間に差は見られなかったこと、アイデア数には差が見られなかったことが示された。また、実験結果と半構造化インタビューで得られた意見を元に、提案手法の効果と制約を検証し、今後の改善指針を明らかにした。実験2では、提案手法と、より優れた表示方法と考えられるカーソル位置にビデオを表示する条件の2条件で比較実験を行った。結果として、会議に対する満足度は提案手法よりもビデオの方が高く、アイデア数には差が見られなかったことが示された。また、実験結果と実験後アンケートおよび半構造化インタビューで得られた意見を元に、提案手法の効果と制約を検証し、さらなる今後の改善指針を明らかにした。

目次

第1章	はじめに	1
1.1	会議の種類	1
1.2	オンラインホワイトボードとは	1
1.3	オンラインホワイトボードの利用における課題	1
1.4	目的とアプローチ	2
1.5	貢献	2
1.6	本論文の構成	2
第2章	関連研究	4
2.1	オンラインコンテンツの共同編集における非言語情報伝達	4
2.2	オンライン会議における非言語情報の提示手法	4
2.3	本研究の位置づけ	5
第3章	設計指針	6
3.1	リアクションの推定手法	6
3.2	リアクションの表示手法	6
第4章	実験1	8
4.1	実装	8
4.1.1	カーソル位置の取得と表示	8
4.1.2	リアクションの取得と表示	8
4.2	実験	10
4.2.1	実験条件	10
4.2.2	仮説	11
4.2.3	実験参加者	11
4.2.4	実験タスク	12
4.2.5	実験環境	12
4.2.6	評価指標	13
4.3	結果および考察	13
4.3.1	リアクション数	14
	既存手法条件および提案手法条件で表示されたリアクション数	14
	各条件中に自動検出されたリアクション数	14

	各条件で自動検出されたリアクションおよび既存手法条件で押された リアクションの種類と数	15
4.3.2	会議に対する評価を測るアンケート	16
	満足度	16
	一体感	16
	リーダーシップ	16
	情報処理	16
4.3.3	アイデア数	17
4.3.4	各条件の好みに関するアンケート	18
4.3.5	半構造化インタビュー	19
	既存手法と提案手法の違い	19
	提案手法とビデオの違い	19
	非言語情報の表示位置	19
	リアクションの表示方法	19
	リアクションの種類	20
4.3.6	追加アンケート, 追加実験	20
	追加アンケート	20
	追加実験	22
4.3.7	考察	23
4.4	議論	25
4.4.1	本実験の制約, 今後の課題	25
	自動検出の認識精度	25
	リアクションの判別が困難	25
	リアクションの種類の設定	25
	誰のリアクションであるか分かりづらい	26
4.4.2	今後の展望	26
	ボタンと自動取得の併用, 使い分け	26
第 5 章	実験 2	27
5.1	実装	27
5.1.1	ホワイトボードの機能	27
5.1.2	映像, 音声の取得と表示	28
5.1.3	リアクションの取得と表示	28
5.1.4	他の参加者との同期	29
5.2	実験	30
5.2.1	実験条件	30
5.2.2	仮説	31
5.2.3	実験参加者	31
5.2.4	実験タスク	32

5.2.5	実験環境	32
5.2.6	評価指標	32
5.3	結果および考察	33
5.3.1	リアクション数	33
	各条件中に自動検出されたリアクション数	33
	各条件で自動検出されたリアクションおよび既存手法条件で押された リアクションの種類と数	34
5.3.2	会議に対する評価を測るアンケート	35
	満足度	35
	一体感	35
	リーダーシップ	35
	情報処理	35
5.3.3	アイデア数	36
5.3.4	実験後アンケート	37
	条件の好みについて	37
	入力したいと思ったが、入力できなかったリアクション	37
	入力したいと思ってなかったが、入力されたリアクション	38
	実際に使用したかどうかに関わらず、会議相手に伝えたいと思う機会 があったリアクション	39
5.3.5	半構造化インタビュー	39
	絵文字およびビデオのサイズについて	39
	会議の前後半での違い	40
	リアクションの種類	40
5.3.6	考察	40
5.4	議論	41
5.4.1	本実験の制約、今後の課題	41
	自動検出の認識精度	41
	実験に用いたモニタの画面サイズ	42
	実験参加者間の関係	42
5.4.2	今後の展望	42
	リアクションの種類の設定	42
	絵文字およびビデオの表示サイズ	42
第 6 章	おわりに	43
	謝辞	44
	参考文献	45
付 録 A	実験におけるアンケート	49

A.1	タスク後アンケート	49
A.2	追加アンケート	54
A.3	実験後アンケート	56

図目次

3.1	リアクション推定および表示の流れ.	7
4.1	表示されるリアクションの一覧.	9
4.2	提案手法を用いた実験時のスクリーンショット. 提案手法部を抜粋したものを合成している.	10
4.3	各条件の概略図.	11
4.4	実験環境.	13
4.5	既存手法条件および提案手法条件で表示されたリアクション数. 有意差が出た条件の間に*を示す.	14
4.6	各条件中に自動検出されたリアクション数.	15
4.7	各条件で自動検出されたリアクションおよび既存手法条件で押されたリアクションの種類と数.	15
4.8	アンケート結果. 有意差が出た条件の間に*を示す.	17
4.9	オンラインホワイトボード上に残されたアイデア数.	17
4.10	好みの順位付けアンケート結果.	18
4.11	入力したいと思ったが, 入力できなかったリアクション.	21
4.12	入力したいと思ってなかったが, 入力されたリアクション.	21
4.13	実際に使用したかどうかに関わらず, 会議相手に伝えたいと思う機会があったリアクション.	22
4.14	正しくリアクションをとれていたフレームの割合.	23
5.1	表示される選択メニュー.	28
5.2	表示されるリアクションの一覧.	30
5.3	各条件の概略図.	31
5.4	各条件中に自動検出されたリアクション数.	34
5.5	各条件で自動検出されたリアクションおよび既存手法条件で押されたリアクションの種類と数.	34
5.6	アンケート結果. 有意差が出た条件の間に*を示す.	36
5.7	オンラインホワイトボード上に残されたアイデア数.	36
5.8	条件の好みについてのアンケート結果.	37
5.9	入力したいと思ったが, 入力できなかったリアクション.	38
5.10	入力したいと思ってなかったが, 入力されたリアクション.	38

5.11 実際に使用したかどうかに関わらず，会議相手に伝えたいと思う機会があった リアクション.	39
---	----

表 目 次

4.1 実験 1 の実験順序.	12
5.1 実験 2 の実験順序.	32

第1章 はじめに

新型コロナウイルスの流行に伴い、リモートワークが普及し、オンライン会議の需要が高まっている。現在主流となっている Zoom に代表されるオンライン会議ツール [Mic,Zoo,Gooc] では、ビデオ通話によりコミュニケーションをとる形態が一般的である。ビデオ通話では、カメラによって撮影された参加者の映像が共有されるため、表情やジェスチャなどの非言語情報を共有することができる。一方で、会議中に何らかのコンテンツが共有され、そちらを見る必要がある場合には、ビデオ通話画面を見ることが困難となる場合がある。そのような状況として、本研究ではオンラインホワイトボードを用いたオンライン会議に着目する。

1.1 会議の種類

会議は伝達会議、調整会議、決定会議および創造会議の4種類に分類されている [会議 08]。伝達会議とは、情報の伝達および確認を行う会議である。調整会議とは、予算、人員、日程等の調整を行う会議である。決定会議とは、物事、方針の決定を行う会議である。創造会議とは、課題に対するアイデアを出し、問題解決を図る会議である。これらの会議のうち特に創造会議は、会議参加者各人がアイデアを出し合い、それを整理する必要がある性質上、ホワイトボードとの親和性が高い。そのため、オンラインで創造会議が行われる際はしばしばオンラインホワイトボードが用いられている。

1.2 オンラインホワイトボードとは

Googole Jamboard, Miro をはじめとするオンラインホワイトボード [Goob,mir,MUR,Fig,Pit] は、ボードが複数人の参加者間でオンライン上で共有され、リアルタイムに共同編集を行うことが可能な仮想的なホワイトボードである。一般的に、テキストや図形、手書きなどを通して情報共有により参加者間の議論を支援する。また、ボード上に各参加者のカーソル位置を表示することにより、対面状況での指差し等のジェスチャの代替を可能としている。

1.3 オンラインホワイトボードの利用における課題

オンラインホワイトボードを用いた創造会議では、ボード上のコンテンツを見ながら議論を行うため、オンラインホワイトボードをディスプレイの最前面に表示し続ける必要がある。

そのため、ラップトップを使用する時など、表示領域の制約がある場合、オンライン会議ツールがバックグラウンドで実行されビデオ映像を見ることができない場合がある。本研究では、このような表示領域の制約がある状況を研究対象とする。

そのような場面で非言語情報を共有するため、既存のオンラインホワイトボードには、ホワイトボード上のボタンにより絵文字を用いたリアクションを伝達する機能が実装されている。ボタンにより伝達されるリアクションは、ボードの中心から離れた領域に一時的に重畳表示される。例えば Miro では画面下端から湧き上がるように表示される。ここでの絵文字のサイズはコンテンツに対して小さく、表示時間も短いため作業の妨げになりにくい。ただし、ツール上に配置されたボタンを能動的に押す必要があるため、表情やジェスチャにより非言語情報が常時共有されるビデオ通話と比較して、伝達される情報量が制限される。また、ボードの中心から離れた領域に表示されることから、リアクションを見るために視線の移動が生じる。

1.4 目的とアプローチ

本研究の目的は、オンラインホワイトボードを用いたオンライン会議において、ホワイトボード上での作業を妨げることなく会議参加者の非言語情報を伝達することである。そのために、カメラを用いて非言語情報を自動取得し、リアクションを推定して各ユーザのマウスカーソル横に絵文字として提示するシステムを提案する。ここで用いる非言語情報は表情、頭部姿勢、手の姿勢、位置を用いる。

1.5 貢献

本研究の貢献は以下の通りである。

- オンラインホワイトボードを用いるオンライン会議において、小さな表示領域で非言語情報を伝達する手法を示した。
- 提案手法を用いて実験を行い、本手法が会議に与える影響を調査し、リアクションボタンを用いる手法と比較して会議に対する満足度の向上が見られた。
- システムを更新し、カーソルにビデオが追従する手法と比較する 2 回目の実験を行った。

1.6 本論文の構成

本論文の構成は以下の通りである。第 1 章においては、本研究の背景、本研究で扱う会議およびオンラインホワイトボードの説明、オンラインホワイトボードの課題、目的とアプローチ、および貢献を示す。第 2 章においては、本研究に関連する研究を述べ、本研究の位置付けを示す。第 3 章においては、本研究にて検討する手法の設計指針について述べる。第 4 章に

おいては，提案手法の効果を検証するための1つ目の実験に関する，実験用アプリケーションの実装，実験内容，結果，考察および議論について述べる．第5章においては，実験1の結果を基に改善したシステムの実装と，その評価実験の内容，結果，考察および議論について述べる．第6章においては，本研究の結論を述べる．

第2章 関連研究

本章では、まず、本研究の対象となるオンラインホワイトボードの共同編集における非言語情報伝達に関連する研究事例として、オンラインコンテンツの共同編集における非言語情報伝達に関する研究について述べる。次に、本研究における提案手法の設計指針検討のために調査したオンライン会議における非言語情報の提示手法、および非言語情報の自動取得に関する先行研究について説明した後、本研究の方針について述べる。

2.1 オンラインコンテンツの共同編集における非言語情報伝達

小俣ら [小俣 09] は、遠隔地間でオンラインコンテンツ（協調パズルシステム）を共同編集する際に、カーソル位置およびユーザの映像を影化したものを共有する“仮想影カーソル”を提案した。仮想影カーソルでは、身振りや手振りにより非言語情報の伝達が可能である。また、カーソル位置に影が追従する設計により、ビデオ映像を別のウィンドウで開く必要がなく、オンラインコンテンツの編集と非言語情報の取得をシームレスに行うことが可能である。ただし、影を用いることにより顔の表情を伝達することが不可能に近く、相手の感情を判断することが困難となる。コミュニケーションの中で表情は、相手の感情を判断する際に大きく影響することが Goodenough ら [GT31] および Ekman ら [EFE82] の研究により示されている。また、非言語情報を伝達する上ではある程度のサイズ感で影を提示する必要がある、コンテンツに重畳される領域が大きくなる。よって本研究では、カーソル位置に非言語情報を提示する方法を踏襲しつつ、コンテンツと重複する領域を小さくするため非言語情報の提示に絵文字を用いるアプローチを採用する。

2.2 オンライン会議における非言語情報の提示手法

オンライン会議における非言語情報提示手法としてはビデオを用いる手法が一般的であり [Mic, Zoo, Gooc], その効果について多くの研究が行われている [DJMW98, OOM95, 岸本 22]。ビデオの代わりにアバタを用いる機能も実用化されており [Mic, Zoo], その効果についても研究が行われている [BYMS06]。ただし、本研究は表示領域の制約がありビデオ映像を見ることができない状況を研究対象とすることから、ビデオやアバタ映像以外で非言語情報の伝達を行う。

ビデオ通話やアバタ映像以外で非言語情報の伝達を行う手法として、既存のビデオ会議ツール [Mic, Zoo, Gooc] やオンラインホワイトボード [Goob, mir, MUR, Fig, Pit] ではツール上のボ

タンにより絵文字を用いたリアクションをコンテンツに重畳して提示する機能が実装されている。絵文字をコンテンツ上に重畳表示する手法は、ビデオや影カーソル [小俣 09] のように大きな表示領域を必要とせず、作業の妨げになりにくい。一方で、これらの絵文字を使用する手法はツール上に配置されたボタンを能動的に押す必要があるため、非言語情報が常時共有されるビデオ通話と比較して、伝達される情報量が制限される。よって、ボタンを能動的に押すのではなく、ユーザーの非言語情報を自動で取得する手法が有効であると考えた。

非言語情報を自動取得する手法として、Zoom や Teams ではビデオ映像を元に画像認識によりユーザのハンドジェスチャを認識し、対応した絵文字をユーザに提示する機能が実装されている。長尾ら [長尾 21] の研究では、ビデオ映像を元に画像認識によりユーザの頭を縦に振る動きと横に振る動きを検出し、対応した表情をしたアバタを他のユーザに提示した。佐藤ら [佐藤 22] の研究では、同様にユーザの顔を縦に振る動きを検出し、“なるほど” の文字を映像中の顔周辺に表示している。Namikawa ら [NSI⁺21] の研究では、ビデオ映像を元に画像認識により表情を検出し、ビデオ映像の顔の位置に、表情と対応する絵文字を重畳表示する手法を提案している。Karras ら [KAL⁺17] の研究では、ユーザの発話時の音声入力から感情推定を行い、その感情を 3D アバタに反映させ、他のユーザに提示した。Oomori ら [OSM⁺20] の研究では、難聴者のために、音声から感情推定を行い音声を文字起こししたテキストとともに絵文字を表示している。このように、非言語情報の自動取得方法としてはビデオ映像を元にした表情・動作認識と、音声を用いた感情推定が用いられる。本研究において想定する絵文字を用いたリアクションでは、感情を表す表情だけでなく、ハンドジェスチャ（サムズアップ等）のような表象的な絵文字も伝達することが期待されることから、音声でなくビデオ映像を元に非言語情報を取得する手法が有効であると考えた。

2.3 本研究の位置づけ

関連研究より、オンライン会議における他の参加者の非言語情報の提示は重要視されており、より効果的な提示を行うため様々な方法が模索されている。しかし、オンラインホワイトボード等のオンラインコンテンツの共同編集を行いながらの状況に適応させた非言語情報の提示手法についての研究は数少なく、コンテンツへの重畳に配慮した手法は検討されていない。そこで本研究では、ビデオから参加者のリアクションを推定し、絵文字として他の参加者に提示する手法を提案する。これにより、ビデオを表示するよりも小さな表示領域で、ビデオに近い量の非言語情報伝達が可能になると考えている。

第3章 設計指針

本研究では、オンラインホワイトボードを用いるオンライン会議において、ホワイトボードの作業を妨げず、作業と同時に他の参加者の非言語情報を共有することを目的とする。そのために、ビデオ映像から表情、頭部、手の動作によるリアクション推定を行い、推定結果を絵文字として表示する手法を提案する。これにより、ビデオよりも表示領域を小さくし、ホワイトボードの作業を妨げることなく非言語情報を共有できると考えている。本章では、提案手法の設計指針として、リアクションの推定手法および表示手法について述べる。なお、後述する第4章実験1と第5章実験2ではシステムの実装内容が異なるため、本章では共通する設計指針のみを説明し、具体的な実装内容は各章内で説明を行う。

3.1 リアクションの推定手法

提案手法において、参加者の非言語情報を利用してリアクションを推定する。リアクションは、顔映像から取得できる表情、頭部姿勢、および映像に映る手から推定を行う。また、表情、頭部姿勢、手の姿勢および位置の推定に関しては、外部の推定ツールを用いて行う。外部の推定ツールにて直接推定できないリアクションは、閾値処理をすることにより推定を行う。

3.2 リアクションの表示手法

推定されたリアクションは、絵文字化を行い、個人と結びつけられたカーソル位置に表示を行う。これは、従来のボタンを押した際のリアクションの表示方法である、画面端に表示を行う、または画面端から浮かび上がるように表示を行う方法と比較し、リアクションを行った個人を特定することが容易になること、画面中央部での表示が可能となりホワイトボード上での作業と両立できることが考えられる。提案手法は従来のボタンを押すリアクションの代替手法として考えているため、参加者がリアクションを取っていないとき（無表情であるとき等）には、何も表示を行わない。また、各リアクションは gif 画像を用いて表示を行う。リアクションの推定および表示の流れを図 3.1 に示す。

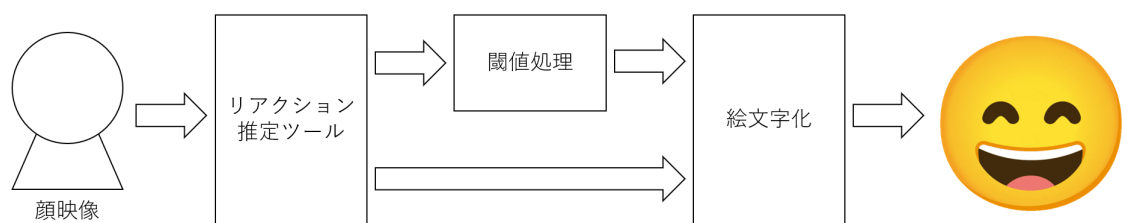


図 3.1: リアクション推定および表示の流れ.

第4章 実験1

提案手法である，顔映像からリアクションを推定し，絵文字化してカーソル位置に表示する手法の効果を検証するため，実験を行った．本実験は，オンラインホワイトボードを用いたオンライン会議で一般的に用いられている，ビデオ共有をせずオンラインホワイトボードのリアクションボタン機能のみを用いる条件および，ビデオ共有を行い画面端に表示する条件の2条件との比較実験とした．本章では，まず本実験に用いた実験用アプリケーションの実装について述べたのち，実験内容，結果，考察，議論について述べる．

4.1 実装

本システムはC++のWindowsAPIを用いて，Windowsアプリケーションとして実装を行った．このアプリケーション単体ではオンライン会議を行うことができず，既存のオンラインホワイトボードツールおよび通話ツールと併用して用いる．

アプリケーションは，他のユーザのマウスカーソル，ユーザ名およびリアクションのみを表示し他の領域を透過するオーバーレイウィンドウ，会議相手を登録するユーザ登録ウィンドウの2つのウィンドウを生成する．ユーザ登録ウィンドウでは，会議を始める際に，他のユーザのユーザ名とIPアドレスの登録を行う．各ユーザのアプリケーション同士はP2PのUDP通信で，カーソル位置およびリアクションのデータをやり取りしている．

4.1.1 カーソル位置の取得と表示

カーソル位置に関しては，WinAPIの関数からディスプレイのサイズと現在のカーソル位置を取得し，ディスプレイサイズとの比率で送信する．オーバーレイウィンドウに，他ユーザから受信したデータをもとに，各ユーザごとに別の色（今回の実験では赤，緑，青色）の矢印型アイコンを表示する．

4.1.2 リアクションの取得と表示

リアクションに関しては，最初にOpenFace [BZLM18]を用いて，ウェブカメラを用いて取得した顔画像から，Facial Action Coding System [EF78]のAction Unitおよび頭部の回転の値を推定，出力させる．Facial Action Coding Systemとは，表情筋の動きを基に表情を分類するものである．Action Unitとは，各表情筋が動いているか否かを示すものであり，これにより

表情および感情推定が行えるようになっている。OpenFaceでは、Action Unitを真偽値で出力させることができるため、Action Unitの組み合わせから基本6感情（喜び、悲しみ、驚き、恐れ、怒り、嫌悪）の推定を行った。具体的には喜びは6（頬を持ち上げる）+12（唇両端を引き上げる）、悲しみは1（眉の内側を上げる）+4（眉を下げる）+15（唇両端を下げる）、驚きは1（眉の内側を上げる）+2（眉の外側を上げる）+5（上瞼を上げる）+26（顎を下げて唇を開く）、恐れは1（眉の内側を上げる）+2（眉の外側を上げる）+4（眉を下げる）+5（上瞼を上げる）+7（瞼を緊張させる）+20（唇両端を横に引く）+26（顎を下げて唇を開く）、怒りは4（眉を下げる）+5（上瞼を上げる）+7（瞼を緊張させる）+23（唇を固く閉じる）、嫌悪は9（鼻にしわを寄せる）+15（唇両端を下げる）+17（顎を上げる）のそれぞれのAction Unitが、OpenFaceでの出力すべてが真となった時に感情が表出していると判定した。頭部の回転からは、うなずきおよび思考中のジェスチャを認識する。うなずきは、過去0.5秒以内に頭部の角度のうち垂直方向に1秒あたり30度動く速度での回転が行われている最中に動く方向が変化したら（下向きの動きが上向きに切り替わるなど）うなずいていると判定した。また、思考中は画面外に顔を向けたら考え事をしており、頭部回転により顔が画面外に向いているときに思考中を表す思考中のジェスチャと判定した。基本6感情（喜び、悲しみ、驚き、恐れ、怒り、嫌悪）、うなずきおよび思考中に、リアクションを取っていないを加えた9種類をユーザのステータスとして登録し、他のすべてのユーザに送信する。オーバーレイウィンドウに、他のユーザから受信したデータをもとに、リアクションを描画する。リアクションはユーザのステータスをもとに絵文字のgif画像に変換し、カーソル位置の横に、ユーザ名とともに表示する。また、ユーザのステータスがリアクションを取っていない場合は何も表示しない。表示されるリアクションの一覧を図4.1に示す。

また、自身が取ったリアクションをタイムスタンプとともにログに残し、実験結果として用いる。提案手法を用いた実験中の画面全体と提案手法部を抜粋したものを図4.2に示す。

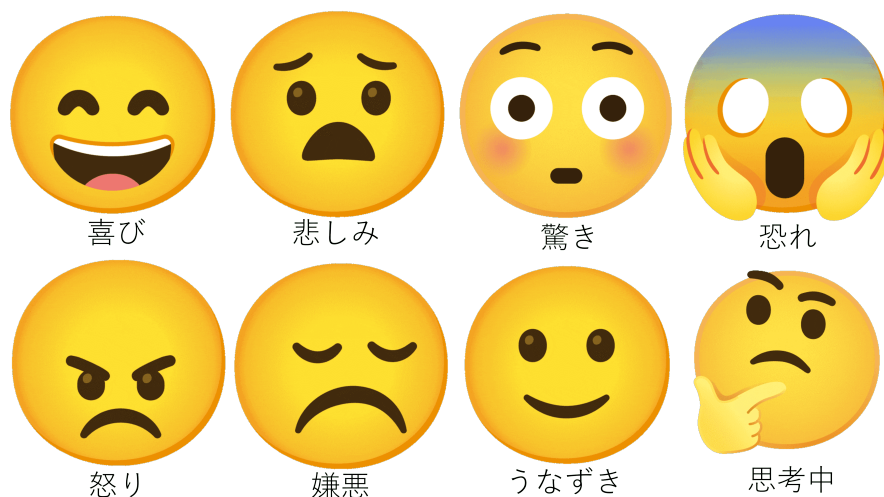


図 4.1: 表示されるリアクションの一覧。



図 4.2: 提案手法を用いた実験時のスクリーンショット。提案手法部を抜粋したものを合成している。

4.2 実験

提案手法の効果を明らかにするために、実装したアプリケーションを用いた実験を行った。

4.2.1 実験条件

実験条件は以下の 3 条件で、すべて参加者内配置で実験を行う。

- 提案手法条件
- 既存手法条件
- ビデオ条件

すべての条件においてオンラインホワイトボードには google jamboard [Goob], 通話およびビデオ共有ツールには google meets [Gooc] を用いて会議を行う。

提案手法条件では、実装したアプリケーションを用いて会議を行う。リアクションは自動取得され、絵文字としてカーソル位置に表示される。

既存手法条件では、提案手法と同様のカーソル共有に加え、ボタンを押すことによりリアクションを共有することができる。ボタンは画面右上 1 列に、提案手法と同じリアクション 8 種類が配置されており、共有されたリアクションはユーザ名とともに画面下部から浮かび上がるように表示される。また、ユーザの表情推移を分析するために、提案手法と同様に顔映像からリアクションの自動取得を行い、表示はせず押されたボタンの情報とともにログに残す。

ビデオ条件では、提案手法と同様のカーソル共有に加え、google jamboard 上に google meets を表示する機能を用い、画面右端に他のユーザのビデオを表示しながら会議を行う。ビデオが表示される領域分、オンラインホワイトボードの表示領域が小さくなる。また、ユーザの表情推移を分析するために、提案手法と同様に顔映像からリアクションの自動取得を行い、表示はせずログに残す。

各条件の概略図を図 4.3 に示す。



図 4.3: 各条件の概略図.

4.2.2 仮説

会議中に他の参加者の非言語情報を知る機会は既存手法よりも提案手法の方が多くなると考えている。また、自動検出していることを参加者に伝えることにより、提案手法においてはユーザが意図的に入力を行おうとし、表情、頭部回転の変化量が増え、自動検出されるリアクション量が他の条件よりも増加することを見込んでいる。他の参加者の非言語情報を知る機会が増えることにより、会議がスムーズに進行されるようになり、ユーザの会議に対する満足度の向上、会議中のアイデア量の増加につながると考えている。また、提案手法により他の参加者の非言語情報を増加させることにより、会議中にユーザが認識する他者の非言語情報が十分量に達し、会議の満足度およびアイデア数がビデオと同等にまで近づくことが期待できる。したがって、以下の仮説を立て、検証の対象とした。

- H1 会議中に表示されるリアクションの数は既存手法よりも提案手法の方が多くなる。
- H2 会議中に自動検出されるリアクションの数は提案手法が最も多くなる。
- H3 会議の満足度は既存手法よりも提案手法の方が高く、提案手法とビデオは同程度になる。
- H4 会議中のアイデア数は既存手法よりも提案手法の方が多く、提案手法とビデオは同程度になる。

4.2.3 実験参加者

実験参加者は計 24 名（うち女性 3 名、平均 22.6 歳、標準偏差 1.35 歳）の大学生および大学院生である。4 人 1 組のグループ分けは、実験前に著者がランダムに行った。実験は参加者

内配置で行われ、各グループが3条件すべてを実施した。実験参加者同士は以前から面識があり、スムーズに会議を行うことができる関係性であった。

4.2.4 実験タスク

最初に、作成したソフトウェア、google meets [Gooc], google jamboard [Goob] の操作に関する説明を行った後、会議中に発言量およびオンラインホワイトボードへの書き込み量が偏らないよう、議長、書記等の明確な役割分担をすることを避けるよう指示した。また、会議中に出的アイデアは google jamboard の付箋機能を用いて書き残すよう指示した。

その後、条件手法の入力の練習、実際の会議、会議に対する評価を測るアンケートへの回答を1セットとし、計3セット行った。会議は7分間であった。また、会議内容は創造会議に関する先行研究 [MP59] を基に、ヘルメット、レンガ、ガムテープについて、新しい使用用途をできる限り多く考えるという内容とした。条件および会議内容の順序は、順序効果に配慮し、ラテン方格でランダム化することで表 4.1 の通りに行った。すべて終了した後に各条件の好みに関するアンケートおよび半構造化インタビューを行った。

グループ番号	会議 1	会議 2	会議 3
1	既存手法-ヘルメット	ビデオ-レンガ	提案手法-ガムテープ
2	提案手法-ヘルメット	既存手法-ガムテープ	ビデオ-レンガ
3	提案手法-レンガ	ビデオ-ヘルメット	既存手法-ガムテープ
4	ビデオ-ヘルメット	既存手法-レンガ	提案手法-ガムテープ
5	ビデオ-ガムテープ	提案手法-レンガ	既存手法-ヘルメット
6	既存手法-レンガ	提案手法-ヘルメット	ビデオ-ガムテープ

表 4.1: 実験 1 の実験順序.

4.2.5 実験環境

互いの音声聞こえないように、別々の部屋に配置されたラップトップコンピュータを用いて実験を行った。ラップトップコンピュータとモニタを接続し、実験参加者はモニタのみを見て実験を行った。モニタの違いにより個人差が出ないように、23.8inch のモニタ (MSI Modern MD241P) に統一した。また、実験中は部屋の中は実験参加者 1 人になるようにし、会議に関係のない情報を与えないようにする。実験中の環境を図 4.4 に示す。



図 4.4: 実験環境.

4.2.6 評価指標

評価指標として、会議中のリアクション数、アイデア数、参加者の会議に対する印象のアンケート評価および各条件に関する主観的意見を用いる。リアクション数は実装したアプリケーションのログデータから、アイデア数はオンラインホワイトボード上に残されたアイデアが書かれた付箋の数から集計する。また、各条件の会議後に、参加者の会議に対する印象の評価を測るアンケートを実施し、主観的な評価も行う。アンケート内容は Post ら [W.M.08] の先行研究で rating scale mental effort (RSME) [ZM89] を基に作成された質問票から不要な項目を削除し、和訳したものを用いた。アンケートは満足度、一体感、リーダーシップ、情報処理の4つの指標で評価を行う。ここで評価を行う情報処理とは、参加者が会議中の情報をどの程度理解できていると感じるか、また、情報を活用して会議に影響を及ぼせたと感じるかどうかを測る指標である。実験の最後には実験参加者の主観的意見を収集するため、各条件の好みに関するアンケートおよび半構造化インタビューを行った。各条件の好みに関するアンケートでは、3つの条件の好みを順位付けさせ、理由とともに回答してもらった。半構造化インタビューでは、“各条件の手法をうまく扱えたと思うかどうか”、“実験中に気になったこと、思ったこと”の2項目を起点に話を広げ、多くの意見を収集するようにした。

4.3 結果および考察

本節では、実験の各評価項目についての結果を示し、考察を行う。正規性の検定にはシャピロウィルク検定、等分散性の検定にはバートレット検定を用いた。正規性、等分散性が認められた場合、二元配置分散分析および多重比較としてチューキークレーマー検定を用いた。

正規性，等分散性が認められなかった場合，フリードマン検定およびウィルコクソンの符号順位検定を用いた．また，検定において p 値が 0.05 未満であることを統計的に有意とみなした．なお，統計解析の実行には R [R C16] を用いた．

4.3.1 リアクション数

実験時，すべての条件中で行ったリアクションの自動検出の結果および既存手法条件時に押されたリアクションのログを取得した．ログをもとに各条件中に自動検出されたリアクション，既存手法条件中に押されたリアクションに関して，その数と種類を比較する．

既存手法条件および提案手法条件で表示されたリアクション数

既存手法条件および提案手法条件で表示された 1 セッション中の参加者個人のリアクション数の累計データを箱ひげ図にしたものを図 4.5 に示す．正規性が認められなかったため，ノンパラメトリック検定を行った．ウィルコクソンの符号順位検定を行った結果，既存手法-提案手法間で有意差 ($p < 0.001$) が見られた．

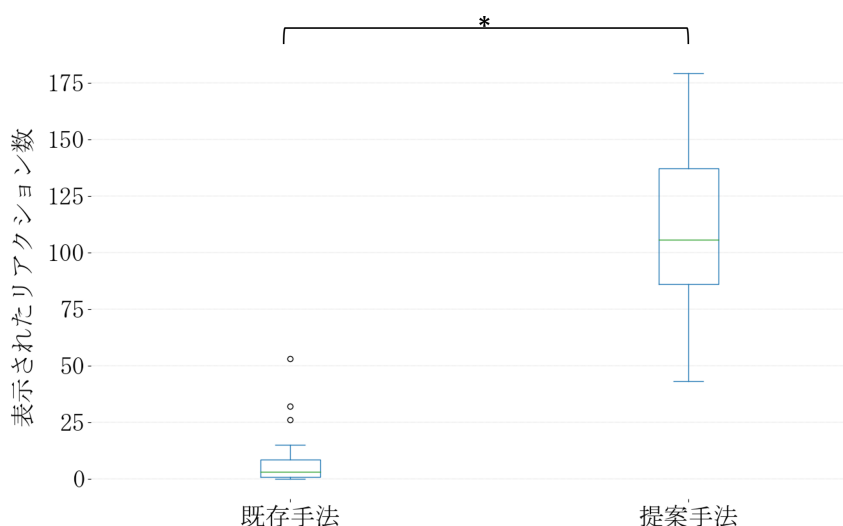


図 4.5: 既存手法条件および提案手法条件で表示されたリアクション数．有意差が出た条件の間に*を示す．

各条件中に自動検出されたリアクション数

1 セッション中に自動検出されたリアクション数のデータの平均を棒グラフにしたものを図 4.6 に示す．エラーバーはデータの標準誤差を示し，以降のグラフも同様である．正規性，等

分散性が認められ、二元配置分散分析を行った結果、手法と議題の交互作用に有意差が見られた ($p=0.012$)。しかし、多重比較を行った結果、すべての条件間で有意差が見られなかった。

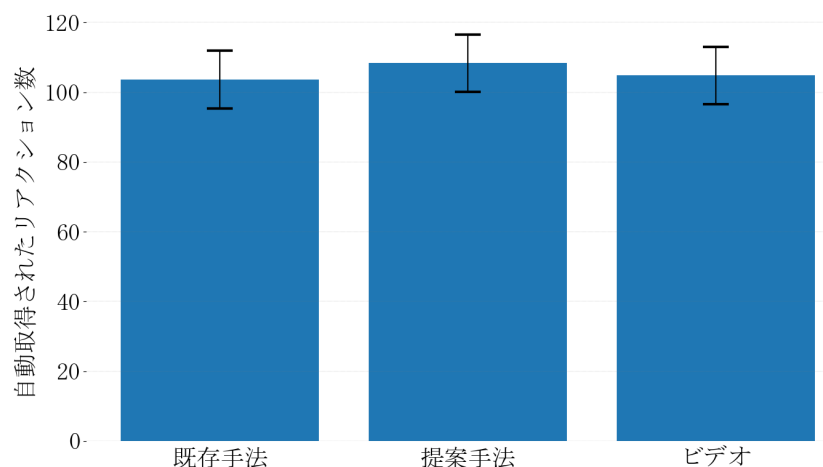


図 4.6: 各条件中に自動検出されたリアクション数.

各条件で自動検出されたリアクションおよび既存手法条件で押されたリアクションの種類と数

各条件で 1 セッション中に自動検出されたリアクションおよび既存手法条件で押されたリアクションの種類と数の累計データを箱ひげ図にしたものを図 4.7 に示す。各条件で自動検出されたリアクションは喜び、悲しみ、思考中、人によっては驚きが多かった。既存手法条件で押されたリアクションは、人によっては喜び、恐れ、思考中が比較的多かった。

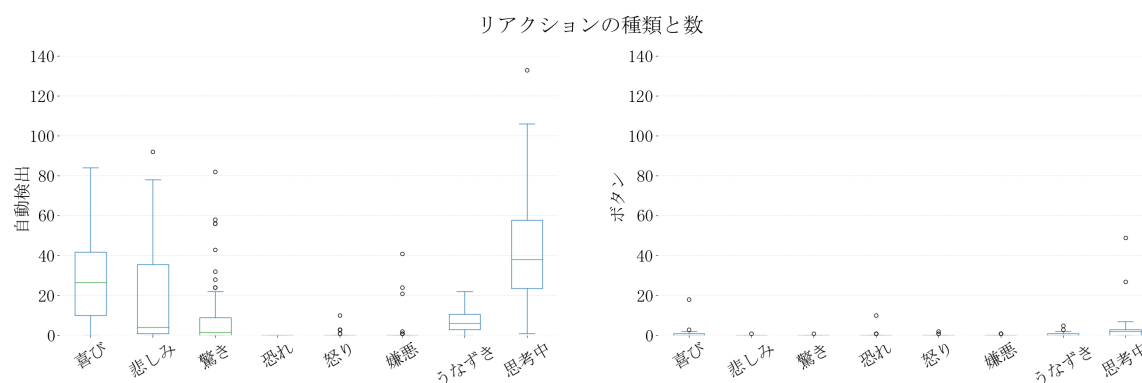


図 4.7: 各条件で自動検出されたリアクションおよび既存手法条件で押されたリアクションの種類と数.

4.3.2 会議に対する評価を測るアンケート

本アンケートは満足度、一体感、リーダーシップ、情報処理の4つの評価指標で会議に対する評価を行う。アンケート結果のデータの平均を棒グラフにしたものを図4.8に示す。すべての評価指標のデータにおいて正規性および等分散性が認められたため、二元配置分散分析を行った。有意差が見られた場合はチューキークレーマー検定により多重比較を行った。

満足度

二元配置分散分析の結果、手法の主効果が有意 ($p=0.011$) であり、議題の主効果および手法と議題の交互作用は有意でなかった。多重比較を行った結果、既存手法-提案手法間 ($p=0.044$) および既存手法-ビデオ間 ($p=0.014$) に有意差が見られた。

一体感

二元配置分散分析の結果、手法の主効果が有意 ($p=0.011$) であり、議題の主効果および手法と議題の交互作用は有意でなかった。多重比較を行った結果、既存手法-ビデオ間 ($p=0.008$) に有意差が見られた。

リーダーシップ

二元配置分散分析の結果、すべての条件間に有意差は見られなかった。

情報処理

本指標は参加者が会議中の情報をどの程度理解できていると感じるか、また、情報を活用して会議に影響を及ぼせたと感じるかどうかを測る指標である。二元配置分散分析の結果、手法の主効果が有意 ($p=0.018$) であり、議題の主効果および手法と議題の交互作用は有意でなかった。多重比較を行った結果、既存手法-ビデオ間 ($p=0.013$) に有意差が見られた。

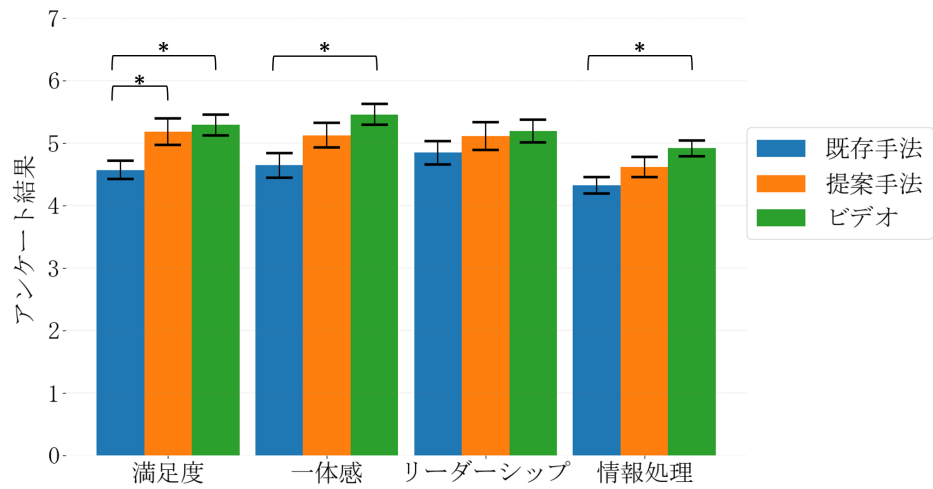


図 4.8: アンケート結果. 有意差が出た条件の間に*を示す.

4.3.3 アイデア数

各セッションにおけるオンラインホワイトボード上に残されたアイデア数の平均データを棒グラフにしたものを図 4.9 に示す. 正規性, 等分散性が認められ, 二元配置分散分析を行った結果, すべての条件間および議題間において有意差は見られなかった.

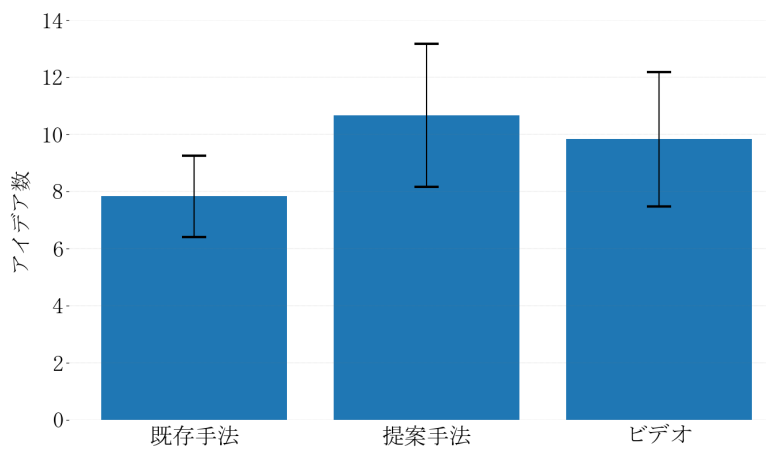


図 4.9: オンラインホワイトボード上に残されたアイデア数.

4.3.4 各条件の好みに関するアンケート

アンケート結果の順位付けのデータを積み上げ棒グラフにしたものを図4.10に示す。フリードマン検定を用いた結果、条件間に有意差が見られた ($p<0.001$)。そのため、ウィルコクソンの符号順位検定を用い、各条件間の比較を行った結果、既存手法-提案手法間 ($p=0.007$)、既存手法-ビデオ間 ($p<0.001$)、提案手法-ビデオ間 ($p=0.021$) で有意差が見られた。

ビデオを好んだ参加者からは、“慣れている”、“相手の顔が見えるため安心感、信頼感がある”等を理由として挙げていた。逆にビデオを好まなかった参加者からは、“自分の表情がずっと見られていることに少し負担を感じた”、“ホワイトボードに集中しビデオをあまり見なかった”という意見が見られた。提案手法を好んだ参加者からは、“操作が必要ないことがよい”、“ユーザの感情がわかりやすい”といった意見が見られた。逆に提案手法を好まなかった参加者からは、“表示されているリアクションが正しいか不安”、“相手のリアクションがあまり目に留まらなかったから”等の意見が見られた。既存手法を好んだ参加者からは、“下から湧き上がる動きがよい”、“会議に参加している感じがした”といった意見が見られた。逆に既存手法を好まなかった参加者からは、“ボタンを押さなかった”、“ボタンを押すタイミングがわからなかった”等の意見が見られた。

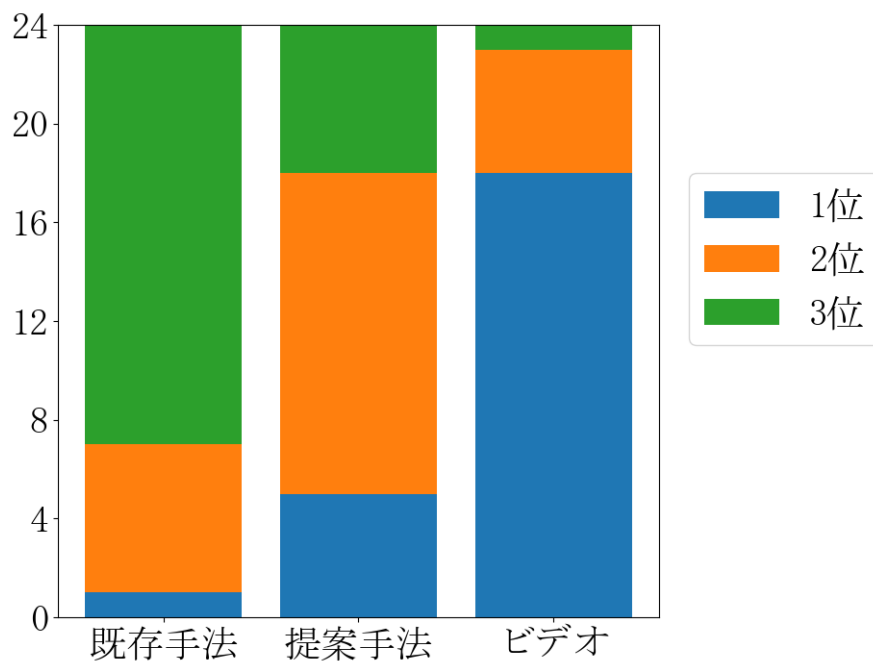


図 4.10: 好みの順位付けアンケート結果.

4.3.5 半構造化インタビュー

既存手法と提案手法の違い

“ボタンによるリアクション共有は聴講者のための機能であり，全員がミュート解除し話し続ける創造会議と相性が悪い”，“ホワイトボードに案を書き込む作業に手を取られ，ボタンを押す余裕がなかった”といった，既存手法に否定的な意見が見られた．一方で，“ボタンの方が表示されるとき動きが大きい目に入りやすく，提案手法よりもリアクションが相手に伝わっていると感じ，安心感があった”といった，表示方法を支持する意見が見られた．また，“ボタンは連打することによってライブ感のようなものが感じられた”といった，既存手法を支持する意見も見られた．

提案手法とビデオの違い

複数人から“ビデオでは相手の顔が直接見えるため安心感，信頼感があった”といった意見が見られた．一方で，“ビデオではずっと顔が出ていて，見られているという意識から疲れる．その点，自動取得でポジティブな面だけ共有してくれれば心理的負担の軽減につながる”といった意見も見られた．

非言語情報の表示位置

“カーソル位置は頻繁に動いてしまうため，誰のカーソル，誰のリアクションであるかを一瞬で判断することが困難だった”といった意見が多く見られた．一方，一部の参加者からは“画面端に表示されたビデオには，目を向ける暇がなく，生かせなかった”という意見も見られた．

リアクションの表示方法

既存手法，提案手法ともに，作業の邪魔になるという意見は見られなかった．一部の参加者からは，“見たいときに見ることができ，見ようと思っていないときには目に入らないためうれしい”等の意見が得られ，好評だった．一部の参加者からは，“リアクションがとられたときにできるだけ見たいから，今のままでは表示が控えめすぎる”，“リアクション発生時に一時的に大きくする，エフェクトをつける等して，強調表示してほしい”といった意見が見られた．また，真顔のときにもカーソル横に真顔の絵文字で表示し続ける方がよいかは，参加者の間で意見が分かれた．ほぼすべての参加者から，リアクションの絵文字が顔文字であり，色，形状が似通っていることから，一目でリアクションを理解できないといった意見が見られた．

リアクションの種類

“ネガティブなリアクション（悲しみ、恐れ、怒り、嫌悪）については、表情を取る機会がないし、相手の参加者がそのリアクションをしていてもどうしようもないし、いらなく感じた”，“思考中は、特に会議中、全員がアイデアを考えていて沈黙状態になったときに表示されていると、他の人も考えているのだなと読み取ることができ、安心感につながった”，“会議として使いやすい挙手やサムズアップなどがあるとよいと思う”といった意見が見られた。

4.3.6 追加アンケート，追加実験

各条件の好みに関するアンケートの自由記述欄および半構造化インタビューで提案手法の認識精度を懸念する意見が見られた。そのため、提案手法のリアクション推定の妥当性を検証する追加アンケートおよび追加実験を行った。

追加アンケート

本実験の参加者 24 名を対象に、実施した本実験中のリアクション推定の精度に関する主観的意見を得るため、追加アンケートを行った。また、実験実施日から 4 ヶ月程度の間が空いている中でのアンケートだったため、実験当時の状況を思い出すために実験時の録画を提供した。以下の質問項目において、該当するリアクションの種類を回答してもらった。

- 入力したいと思ったが、入力できなかったリアクション
- 入力したいと思ってなかったが、入力されたリアクション
- 実際に使用したかどうかに関わらず、会議相手に伝えたいと思う機会があったリアクション

該当した参加者の人数を示す結果はそれぞれ図 4.11、図 4.12、図 4.13 のようになった。

入力したいと思ったが、入力できなかったリアクションについては喜び、驚き、うなずき、思考中が少人数該当した。入力したいと思ってなかったが、入力されたリアクションについては喜び、思考中が 3,4 割程度、悲しみ、驚き、怒り、うなずきが少人数該当した。また、会議相手に伝えたいと思う機会があったリアクションは喜び、驚き、うなずき、思考中が挙げられ、悲しみ、怒りが少人数該当した。

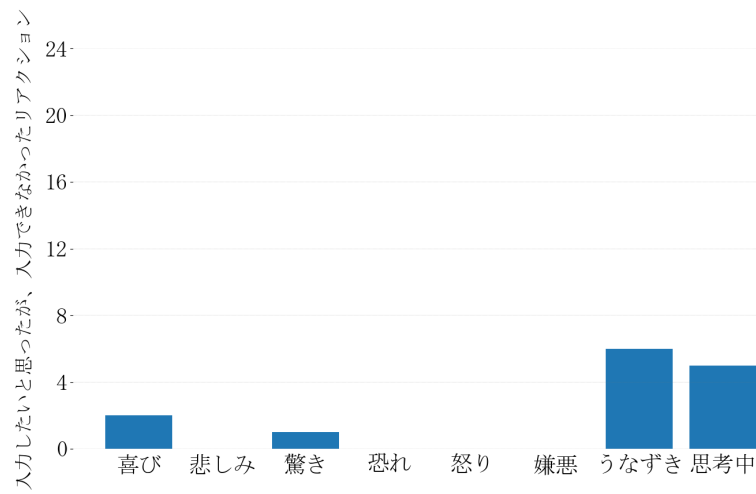


図 4.11: 入力したいと思ったが、入力できなかったリアクション.

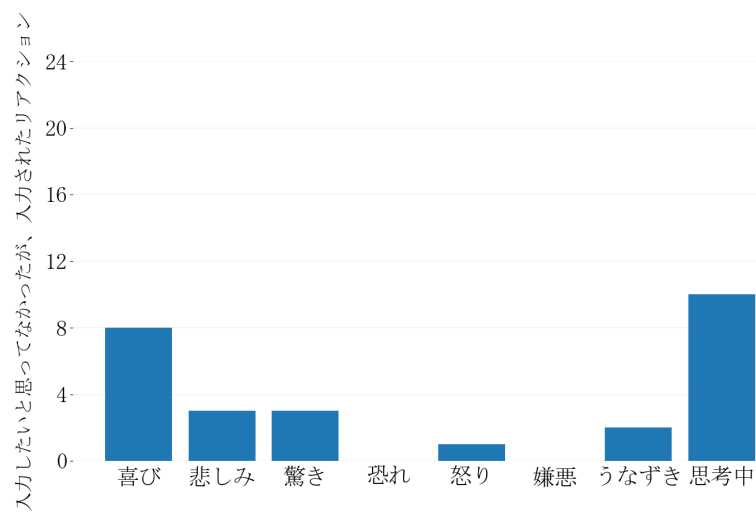


図 4.12: 入力したいと思ってなかったが、入力されたリアクション.

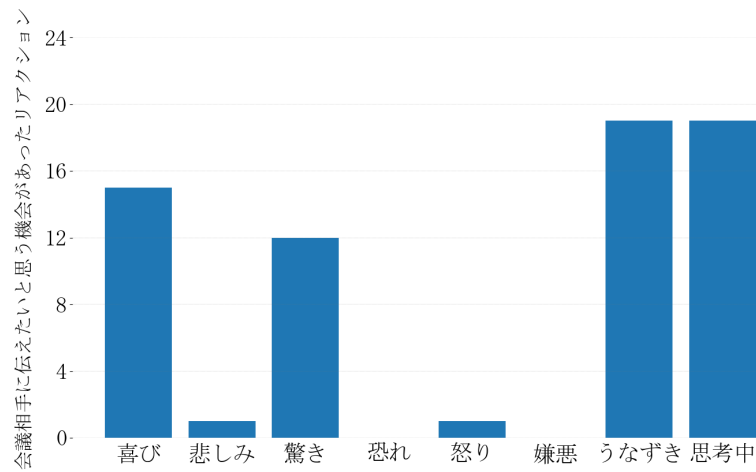


図 4.13: 実際に使用したかどうかに関わらず、会議相手に伝えたいと思う機会があったリアクション。

追加実験

半構造化インタビューや追加アンケートの結果から、本システムでの誤認識が起きている可能性が示唆された。そのため、エラー率を定量的に測定し、本システムを評価する実験を実施した。実験参加者は計 8 名（うち女性 1 名、平均 22.75 歳、標準偏差 0.97 歳）の大学生および大学院生である。実験タスクは、提案手法を用いた条件中、8 種類のリアクションからランダムなものが提示され、表情および頭部回転を変化させそのリアクションをとるというものである。リアクションを指示されてから 6 秒間の準備時間が設けられ、その後 2 秒間でどの程度リアクションを取り続けることができたかを評価する。8 種類のリアクションが 3 回ずつランダムな順序で計 24 回提示されるのを 1 セットとし、5 セット行ってもらった。2 秒間（60 フレーム）のうち、指示されたリアクションと判定されたリアクションを比較し、正しくリアクションをとれていたフレームの割合を各試行ごとに計算した。計算結果から平均を求めた結果は図 4.14 のようになった。

具体的な割合としては喜びが 83.8%、悲しみが 68.2%、驚きが 74.1%、恐れが 0.0%、怒りが 0.0%、嫌悪が 0.0%、うなずきが 71.8%、思考中が 94.1% となった。思考中に関しては 9 割を超える精度、喜び、悲しみ、驚き、うなずきに関しては約 7～8 割程度の精度、恐れ、怒り、嫌悪はリアクションを識別することが困難という結果となった。なお、喜び、驚き、うなずきに関しては実験参加者 8 名全員が指定された表情を出すことができたが、判定区間である 2 秒間ずっと維持し続けることが困難であったことに起因して精度が低下した。また、怒り、嫌悪に関してはそれぞれ 3 名の参加者のみが出すことができる結果となった。

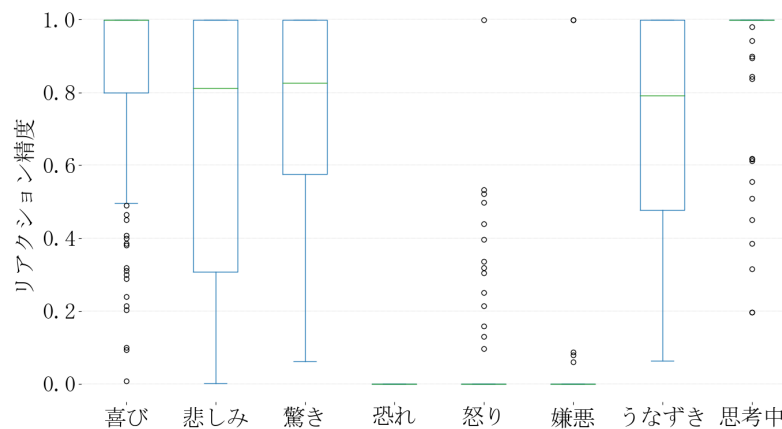


図 4.14: 正しくリアクションをとれていたフレームの割合.

4.3.7 考察

4.3.1 項, 既存手法条件および提案手法条件で表示されたリアクション数より, H1 は支持された. 仮説で述べた通り, 提案手法により共有される非言語情報が増えたことが示されている. それに関連して, 4.3.5 項より, 会議中にボタンを押す余裕がなかったという意見が見られた. これは, 既存手法条件では潜在的にリアクションを共有したい欲求がありつつもそれを共有できていない状態にあった可能性があり, 提案手法条件ではボタンを押す必要がなかったため共有されるリアクションが増えたと考えられる.

4.3.1 項, 各条件中に自動検出されたリアクション数より, H2 は支持されなかった. 仮説ではユーザに自動検出であることを伝えることでユーザが意図的に入力を行おうとし, 提案手法条件時に自動検出されるリアクション数が増加すると考えていた. 原因として, 本システムで自動検出するリアクションがすべて無意識で出るものだったため, ユーザが意図的にリアクションを増やす行動をとらなかったこと, 常にアイデアを考え続ける必要があるため意識を割くことができなかったことが考えられる. 半構造化インタビューで得られた意見の中にも, “意識的に増やそうとはしなかった”, “増やすつもりはあったが, 会議が忙しくできなかった” 等が見られた.

4.3.2 項, 満足度より, H3 は支持された. 仮説で述べた通り, 既存手法よりも提案手法およびビデオの方が会議中に表示されるリアクションの数が増えることにより, 会議の満足度が向上したと考えられる. また, ビデオを用いることなく会議の満足度を向上させることができたことにより, ビデオを用いることができない, または用いたくない状況で有用であると考えられる.

4.3.3 項より, H4 は支持されなかった. 仮説では提案手法により他の参加者の非言語情報を知る機会が増え, 会議がスムーズに進行されるようになり, 会議中のアイデア量が増加す

と考えていた。理由としては、会議の内容が新しい使用用途をできる限り多く考えるというものであり、参加者同士での合意形成等、話を合わせる時間が存在しない。そのため、進行のスムーズさはアイデア量に大きく影響しないからであると考えている。

4.3.1 項、各条件で自動検出されたリアクションおよび既存手法条件で押されたリアクションの種類と数より、リアクションの種類に関して、共通で挙げられている喜び、思考中のリアクションは、ボタンを押すほどユーザが共有したいと考えているリアクションであり、その回数を自動検出により増やすことができたと考えている。また、悲しみおよび驚きのリアクションは、半構造化インタビューで一部の参加者から誤認識されたという意見が見られ、改善すべき点である。一方、既存手法でのみ挙げられた恐れ of リアクションは、実際に恐れを感じているわけではなく、限定的な会話の流れの中で連打させ表示させることで会話を盛り上げる、提案手法では表現できない使われ方をしていた。また、連打という方法はボタン固有の表現方法であり、将来的に提案手法との使い分け、併用を考える際の重要な要素であると考えられる。

4.3.2 項、一体感より、一体感 is 既存手法-ビデオ間で有意差が見られた。参加者はビデオにより他の参加者の顔が直接見えることで、一体感を感じていると考えている。半構造化インタビューでも、“ビデオで他の参加者の顔が見れると安心する”といった意見がみられた。

4.3.2 項、リーダーシップより、リーダーシップにおいて有意差は見られなかった。本実験で行われた会議は、すべての参加者が対等な立場で意見を出し合うものであり、特定の個人によって主導で行われるものではない。また、実験前に注意事項として、発言量、作業量が偏らないよう、議長、書記等の明確な役割を持つことを禁止していた。そのため、リーダーシップの評価で有意差がみられなかったと考えられる。

4.3.2 項、情報処理より、情報処理は既存手法-ビデオ間で有意差が見られた。ビデオの結果が良い原因としては、全参加者が日常的にオンライン会議を使用しており、慣れていることにより、情報の取得および発信がうまく行えていると考えている。また、既存手法の結果が悪い原因としては、大半の参加者がボタンを押さない、または数回程度しか押さなかったため、会議中に共有される非言語情報が少なく、情報の取得がうまく行えていないこと、また、ボタンを押していないため情報を発信していないと感じ、スコアが悪くなったと考えている。

4.3.6 項、追加実験の結果から、恐れ、怒り、嫌悪のリアクションはユーザが意図して入力することは困難という結果となった。ただし、図 4.13 で示された通り、実際の会議中にユーザはこれら 3 つおよび悲しみのリアクションを取りたいと感じた場面がなかったことから、実験結果自体に大きな影響はないと考える。また、喜びおよび驚きに関して追加実験での精度が悪いのは、求められる顔の動きが大きく、継続して表情を取り続けるのが困難であるからだと考えられる。追加実験中に一度も双方のリアクションを取れない人はいなかった。ただし、図 4.11 から、喜びを出したいときにしかなかった人は 2 名、驚きでは 1 名のみだったため、本実験での影響は大きくないと考えられる。また、表情を維持することが困難なことから、一度認識されたら一定時間表示し続ける等の実装の改善が考えられる。うなずきに関しては、実装段階で設けた頭部の回転速度の閾値や切り返し部分で検出する手法に問題があり、喜び、驚きと同様に継続してリアクションを出し続けることが困難だったと考えら

れる。うなずきの適切な閾値や検出方法については今後の調査を行い、検討していく必要がある。思考中に関しては、追加実験の結果から精度自体は問題ないが、追加アンケートで入力したいときにできなかったという意見が見られ、思考中の画面外を向いていたら検出という実装方法に問題がある可能性がある。画面を向きながら思考をしている場面は多く存在し、そのような状況の時も他の会議参加者に思考中であることを共有したいとユーザが感じていると考えられる。今後、より適切な思考中の認識方法を検討していく必要がある。

4.4 議論

4.4.1 本実験の制約、今後の課題

自動検出の認識精度

半構造化インタビューおよび追加アンケートの中で、一部の参加者から意図的に出すことができないリアクションおよび意図していないリアクションの出力が起きたという意見が見られた。また、参加者の中には認識精度に対する不安を感じるという意見も見られた。表情認識の精度に関して、追加実験の結果から恐れ、怒り、嫌悪の4種類は意図して出せない状態であり、また、うなずきおよび思考中は実装方法に問題があった可能性がある。一方で、いずれの表情も他の表情が誤って検出されることは無かった。ただし、ユーザごとの個人差により認識精度に影響が生じている可能性がある。よって今後はOpenFace [BZLM18] 以外の識別手法を用いる、より適切な閾値処理を行う等精度向上のための更新を行う。

リアクションの判別が困難

本研究の実装で用いた絵文字はすべて同じ色、似た形状をしていたため、一目でどのリアクションが取られたか判断できないという意見が多く見られた。また、提案手法の表示方法は控え目であり、リアクションの変化がわかりづらいといった意見も見られた。既存手法でも、画面上部で作業をしていた際には目に入らず、気づけないという意見も見られた。新たな表示方法の案として、顔文字以外の絵文字（ハート、怒りマーク等）を用いる、リアクション表示時に一時的に大きくする、エフェクトを付与する等が挙げられた。今後リアクションの表示方法に関して、邪魔にならず、自然と目に入る表示方法を調査する必要がある。

リアクションの種類の設定

表情からのリアクションには基本6感情をもとに感情の設定を行ったが、ネガティブな感情（悲しみ、恐れ、怒り、嫌悪）は必要ないという意見が多く見られた。また、既存製品でリアクションとして用いられている挙手、サムズアップ、クラッカーなどが新たなリアクションの案として挙げられた。今回使用していないものも含め、必要とされるリアクションの種類を調査する必要がある。

誰のリアクションであるか分かりづらい

本研究の提案手法では、カーソル位置に追従し自動検出したリアクションを表示させている。その結果、多くの実験参加者から、カーソルが動くことにより、一目で誰のリアクションであるか理解できないという意見が見られた。システムでは自身以外のユーザのカーソルはそれぞれ別の色で表示されていたが、色とユーザが対応するには会議時間が短かったと考えられる。改善案として、カーソル同様にリアクションも色分けする、リアクションを意味を損なわない程度にカスタマイズできるようにする（絵文字の髪型等を変える等）が考えられる。今後詳細を検討し実装を行う。

4.4.2 今後の展望

ボタンと自動取得の併用、使い分け

考察で述べた通り、既存手法であるボタンを押すことによるリアクションは、連打できる、実際には感じていないリアクションを表現できるという点で提案手法によるリアクションと異なる。自動取得によるリアクションはビデオの代替として位置し、ボタンとの併用が可能であると考えている。その際問題となるのが、自動取得およびボタンで用いるリアクションの選定である。現時点では、無意識に取られる表情やジェスチャのうちポジティブなものを自動取得として、意識的でないと出てこない表情やジェスチャをネガティブなものも含めてボタンとして用いることを考えている。ただし、意識的でないと出ない表情やジェスチャを取らせることにより、会議への没入感、関与をより感じられ、満足度の向上につながる可能性もあると考えている。自動取得およびボタンで用いる際に適しているリアクションは今後も引き続き調査が必要である。

第5章 実験2

実験1では、提案手法を一般的に広く用いられている手法と比較を行った。一方で, SpatialChat [Spa] 等の一部の製品では、ビデオ共有の画面をオンラインホワイトボード上の好きな位置に配置することができる。これは、一般的に広く用いられている画面端に表示するよりも、非言語情報伝達の面で優れている可能性があると考えた。そこで、本実験では、提案手法とカーソル位置にビデオ共有の画面を配置する手法の比較実験を行う。また、提案手法について、実験1で見られた問題点は改善を行うとともに、オンラインホワイトボード機能やリアクションの取得・提示機能、およびビデオ通話機能等を備え、単独で想定するインタラクションが完結する web アプリケーションとして提案システムを再実装した。本章では、まず本実験に用いた実験用 web アプリケーションの実装について述べたのち、実験内容、結果、考察、議論について述べる。

5.1 実装

本システムは HTML, CSS, JavaScript および Nodejs を用いて web アプリケーションとして実装を行った。このアプリケーションは映像、音声の共有とオンラインホワイトボード機能をまとめたものであり、単一でオンラインホワイトボードを用いたオンライン会議を行うことができる。

アプリケーションは入室する部屋を選択するロビー画面と、実際に会議を行うルーム画面に分かれている。ロビー画面ではルーム名、ユーザ名、実験で用いる手法名（0を提案手法条件、1をビデオ条件とした）を入力し、join ボタンを押すことでルーム画面に遷移する。ルーム画面では同じルームに接続している人の音声再生され、オンラインホワイトボードを扱うことができる。また、オンラインホワイトボード上ではカーソル位置とユーザ名の共有がされており、提案手法条件ではカーソル位置にリアクション、ビデオ条件ではカーソル位置にビデオが表示される。

5.1.1 ホワイトボードの機能

会議中の意見を残すための機能として、簡易的なホワイトボードを実装した。カーソル共有として、自身のカーソル横および他の参加者のカーソル位置に、長方形とユーザ名が表示される。長方形はユーザごとに青赤黄緑橙の5色から被りがないように配布された色で塗りつぶす。ホワイトボード左上に作成した create node ボタンを押すことで付箋を生成する。付

箋は左クリックを1回することにより選択モードに移行，選択モード中に右クリックをすることで選択メニューを表示する．選択メニューの見た目は図 5.1 に示す．

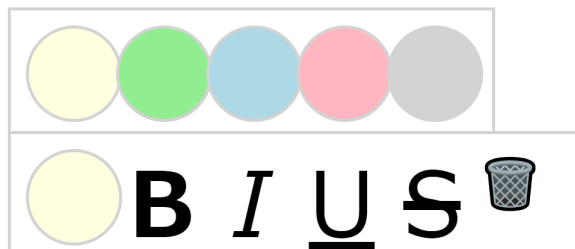


図 5.1: 表示される選択メニュー．

選択メニューの内容は図 5.1 の左から，付箋の色を変更するボタン，付箋内の文字を太字にする（太字を解除する）ボタン，付箋内の文字をイタリックにする（イタリックを解除する）ボタン，付箋内の文字にアンダーラインを付ける（アンダーラインを解除する）ボタン，付箋内の文字に取り消し線を付ける（取り消し線を解除する）ボタン，付箋を削除するボタンの 6 種類である．付箋の色を変更するボタンを押した場合，図 5.1 の上段が展開され，5 色から選択できるボタンが提示される．すでに展開されている場合は非表示にする．また，選択モード時に Ctrl+c を押すことでその付箋をコピーし，Ctrl+v でユーザのカーソル位置にペーストする．付箋を左ダブルクリックすることで付箋内のテキストを編集できる．付箋の中央部分を左クリック長押ししている間，付箋を移動させることができ，付箋の端部分を左クリック長押ししている間，付箋のサイズを変更することができる．また，以上の機能についてすべての変更操作を Ctrl+z で変更前に戻すことができる．

5.1.2 映像，音声の取得と表示

映像，音声の取得と転送，表示は SkyWay [Cor] の JavaScript SDK を用いた．映像および音声は SkyWay 側で提供される TURN サーバを経由してリアルタイムに送受信が行われる．音声は両条件で再生し続けた．映像は提案手法条件では表示せず，ビデオ条件ではカーソル位置の下に 100 × 100px で表示した．

5.1.3 リアクションの取得と表示

リアクションに関しては，face-api.js [jus] の Face Expression Recognition, MediaPipe [Good] の Hand Gesture Recognition および Pose Landmark Detection を用いてリアクションの推定を行う．face-api.js の Face Expression Recognition では，顔映像または画像から顔の特徴点を抽出し，感情推定を行うことができる．MediaPipe の Hand Gesture Recognition では，手の映った映像または画像から，手の姿勢を基にジェスチャ認識を行うことができる．MediaPipe の

Pose Landmark Detection では、人間の映った映像または画像から、全身の特徴点の位置を推定することができる。

リアクションは喜び、驚き、うなずき、思考中、サムズアップ、挙手、拍手の7種類を選定した。表示されるリアクションの一覧を図5.2に示す。喜び、驚き、うなずき、思考中の4種類に関しては、実験1の参加者から、実験1で用いたリアクションのうち悲しみ、恐れ、怒り、嫌悪のネガティブなものは不要という意見が多く見られたため、それ以外の4種類を採用した。サムズアップ、挙手、拍手に関しては、実験1の参加者から得られた追加するとよいと考えられるリアクションのうち、表情、頭部姿勢、手を用いて推定できるものとして採用した。

喜び、驚きは face-api.js の Face Expression Recognition の機能を用いて推定を行った。サムズアップ、挙手は MediaPipe の Hand Gesture Recognition の機能を用いて、ThumpsUp が出力されたらサムズアップ、Palm（手のひらがカメラに向かっている状態）が出力されたら挙手として推定を行った。うなずき、思考中、拍手は MediaPipe の Pose Landmark Detection の機能を用いて推定を行った。うなずきは鼻の特徴点の垂直方向の動きから、0.5 秒間のうちにカメラの画角全体の3%以上の速度で動く方向が切り返されたときにうなずいていると判定し、最後に判定されてから0.5 秒間持続してうなずいていると出力した。思考中は口元に手があるときに思考中と判定した。具体的には、鼻の特徴点と手の中心の特徴点の位置関係が垂直方向に画角全体の10%以内の位置にあるときに思考中と出力した。この際口と手の位置関係にできなかったのは、MediaPipe での出力で、手で口を覆ったときに口の位置の推定が不安定になったためである。拍手は、両手の位置関係が画角全体の10%以内になったときに両手が閉じられているとして、一度両手が閉じられ、離れてから0.5 秒以内にもう一度両手が閉じられた際に拍手をしていると判定し、思考中と同様、最後に拍手をしてから0.5 秒間持続して拍手をしていると出力した。また、口元で拍手を行うと、思考中と拍手が同時に出力されてしまうことを防ぐため、両方が同時に判定された場合は拍手を優先して出力するようにした。

推定の際には、face-api.js および Mediapipe の2つの WebWorker を作成し、ホワイトボード、映像音声を扱っているメインスレッドとは別スレッドにすることにより処理の重い画像処理を行いながらホワイトボード動作、音声、映像が止まらないようにした。WebWorker ではメインスレッドからカメラ画像を受け取り、推定されたリアクションを返却する。メインスレッドからは、最初に1枚画像を入力し、WebWorker からリアクションを受け取り次第、その時点のカメラ画像を再び WebWorker に入力するということを繰り返した。

推定されたリアクションは、メインスレッドで2つの WebWorker の結果を統合し、対応した絵文字の gif 画像を用いてカーソル位置の下部に 50 × 50px で表示した。

5.1.4 他の参加者との同期

ホワイトボードおよびリアクションの内容を参加者間で同期させるため、Google Firebase の Realtime Database [Gooa] を用いた。Realtime Database は json 形式で管理される NoSQL データベースであり、同じ会議内のデータに付箋やユーザの作成、更新が行われたことを受け取る EventListener をセットできる。

データベースには、ルーム名ごとにユーザに関する情報（ユーザ名、割り当てられた色、カーソル位置、SkyWayでの映像のID、現在のリアクション）、ホワイトボード上の付箋の情報（付箋の位置、色、書かれている文章、サイズ、太字イタリック下線取り消し線の有無）、およびリアクションのログ（現在のUNIX時間、現在のリアクション）を格納する。このうち、ユーザ、付箋の追加、削除および各ユーザのリアクション、各付箋に関する情報すべての更新に関してEventListenerをセットし、発火した際にホワイトボード上の情報の更新を行う。また、ユーザはブラウザバック等、会議のページを離れた際にデータベースから削除するようにしている。

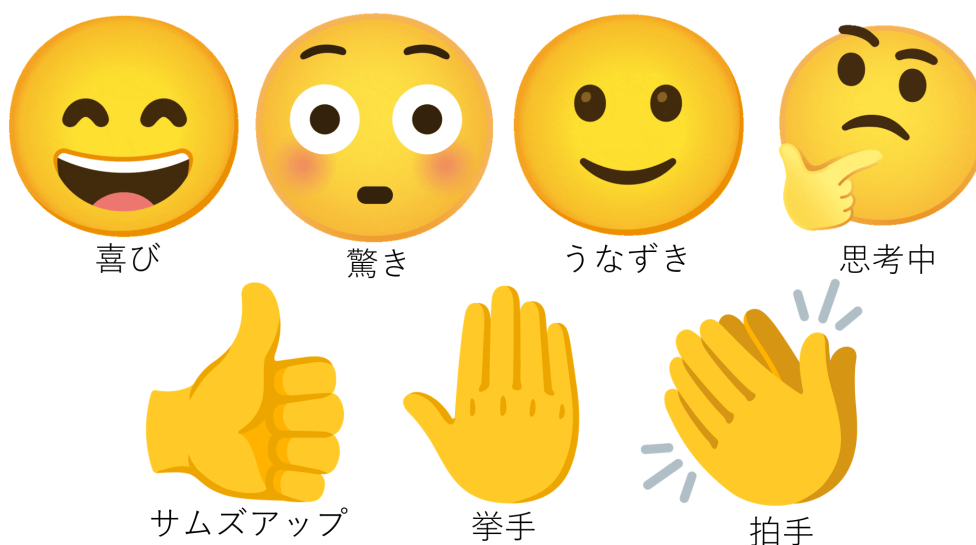


図 5.2: 表示されるリアクションの一覧.

5.2 実験

提案手法の効果を明らかにするために、実装したアプリケーションを用いた実験を行った。

5.2.1 実験条件

実験条件は以下の2条件で、すべて参加者内配置で実験を行う。

- 提案手法条件
- ビデオ条件

両方の条件において、実装した提案システムを用いて会議を行う。

提案手法条件では、ビデオは表示されず、自動取得されたリアクションが絵文字としてホワイトボード上のカーソル位置に表示される。

ビデオ条件では、推定されたリアクションは表示されず、各参加者のカーソル位置に各参加者のビデオを表示しながら会議を行う。また、ユーザの表情推移を分析するために、提案手法と同様に顔映像からリアクションの自動取得を行い、表示はせずログに残す。

各条件の概略図を図 5.3 に示す。

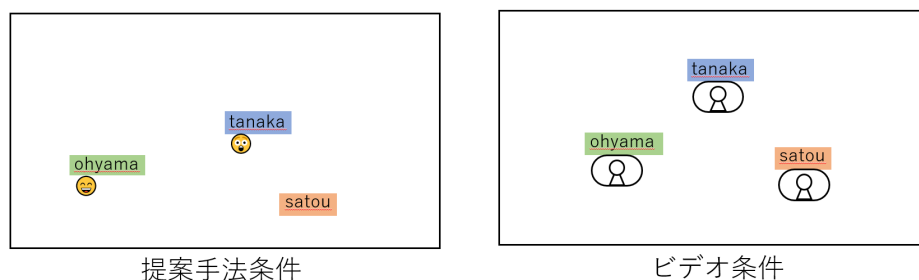


図 5.3: 各条件の概略図.

5.2.2 仮説

リアクションを自動検出していることを参加者に伝えることにより、提案手法においてはユーザが意図的に入力を行おうとし、表情、頭部回転、手の動きの変化量が増え、自動検出されるリアクション量が他の条件よりも増加することを見込んでいる。また、提案手法により他の参加者の非言語情報を増加させることにより、会議中にユーザが認識する他者の非言語情報が十分に伝達され、会議の満足度およびアイデア数がビデオと同等にまで近づくことが期待できると考える。したがって、以下の仮説を立て、検証の対象とした。

H1 会議中に自動検出されるリアクションの数は提案手法の方が多くなる。

H2 会議の満足度は、提案手法とビデオで同程度になる。

H3 会議中のアイデア数は、提案手法とビデオで同程度になる。

5.2.3 実験参加者

実験参加者は計 16 名（うち女性 2 名、平均 22.88 歳、標準偏差 0.78 歳）の大学生および大学院生とした。4 人 1 組のグループ分けは、実験前に著者がランダムに行った。実験は参加者内配置で行われ、各グループが 2 条件両方を実施した。実験参加者同士は以前から面識があり、スムーズに会議を行うことができる関係性であった。

5.2.4 実験タスク

最初に、作成したアプリケーションの操作に関する説明を行った後、会議中に発言量およびオンラインホワイトボードへの書き込み量が偏らないよう、議長、書記等の明確な役割分担をすることを避けるよう指示した。また、会議中に出たアイデアは作成したアプリケーションの付箋機能を用いて書き残すよう指示した。

その後、条件手法の入力の練習、実際の会議、会議に対する評価を測るアンケートへの回答を1セットとし、計2セット行った。会議は前半の7分間と後半の3分間を合わせた計10分間である。会議内容は、前半の7分間で創造会議に関する先行研究 [MP59] を基に、とんかち、調理用ヘラについて、新しい使用用途をできる限り多く考える、後半の3分間で出た新しい使用用途のアイデアの中から一番良いアイデアを決めるという内容とした。会議の前後半の切り替えは、実験分担者から直接実験参加者に指示を行った。条件および会議内容の順序は、順序効果に配慮し表5.1の順序で行った。すべて終了した後に実験後アンケートおよび半構造化インタビューを行った。

グループ番号	会議 1	会議 2
1	提案手法-とんかち	ビデオ-調理用ヘラ
2	提案手法-調理用ヘラ	ビデオ-とんかち
3	ビデオ-とんかち	提案手法-調理用ヘラ
4	ビデオ-調理用ヘラ	提案手法-とんかち

表 5.1: 実験 2 の実験順序.

5.2.5 実験環境

互いの音声が聞こえないように、別々の部屋に配置されたラップトップコンピュータを用いて実験を行った。ラップトップコンピュータとモニタを接続し、実験参加者はモニタのみを見て実験を行った。モニタの違いにより個人差が出ないように、23.8inch のモニタ (MSI Modern MD241P) に統一した。また、実験中は部屋の中は実験参加者1人になるようにし、会議に関係のない情報を与えないようにした。実験中の環境は実験1 (図4.4) と同様とした。

5.2.6 評価指標

評価指標として、会議中のリアクション数、アイデア数、参加者の会議に対する印象のアンケート評価および各条件に関する主観的意見を用いた。リアクション数は実装したアプリケーションのログデータから、アイデア数はオンラインホワイトボード上に残されたアイデアが書かれた付箋の数から集計した。また、各条件の会議後に、参加者の会議に対する印象の評価を測るアンケートを実施し、主観的な評価も行った。アンケート内容は Post ら [W.M.08] の先行研究で rating scale mental effort (RSME) [ZM89] を基に作成された質問票から不要な項

目を削除し、和訳したものを用いた。アンケートは満足度、一体感、リーダーシップ、情報処理の4つの指標で評価を行った。ここで評価を行う情報処理とは、参加者が会議中の情報をどの程度理解できていると感じるか、また、情報を活用して会議に影響を及ぼせたと感じるかどうかを測る指標である。実験の最後には実験参加者の主観的意見を収集するため、実験後アンケートおよび半構造化インタビューを行った。実験後アンケートでは、2つの条件の好みの順位付け、“絵文字手法のとき、入力したいと思ったが、入力できなかったリアクション”、“絵文字手法のとき、入力したいと思ってなかったが、入力されたリアクション”、“手法、実際に使用したかどうかに関わらず、会議相手に伝えたいと思う機会があったリアクション”に回答してもらった。半構造化インタビューでは、“絵文字およびビデオのサイズについて”、“会議の前後半での違い”の2項目を起点に話を広げ、多くの意見を収集するようにした。

5.3 結果および考察

本節では、実験の各評価項目についての結果を示し、考察を行う。正規性の検定にはシャピロウィルク検定、等分散性の検定にはバートレット検定を用いた。正規性、等分散性が認められた場合、二元配置分散分析および多重比較としてチューキークレーマー検定を用いた。正規性、等分散性が認められなかった場合、フリードマン検定およびウィルコクソンの符号順位検定を用いた。また、検定においてp値が0.05未満であることを統計的に有意とみなした。なお、統計解析の実行にはR [R C16]を用いた。

5.3.1 リアクション数

実験時、すべての条件中で行ったリアクションの自動検出の結果および既存手法条件時に押されたリアクションのログを取得した。ログをもとに各条件中に自動検出されたリアクション、既存手法条件中に押されたリアクションに関して、その数と種類を比較する。

各条件中に自動検出されたリアクション数

1セッション中に自動検出されたリアクション数のデータの平均を棒グラフにしたものを図5.4に示す。正規性が認められなかったため、ノンパラメトリック検定を行った。ウィルコクソンの符号順位検定を行った結果、手法間および議題間では有意差は見られなかった。

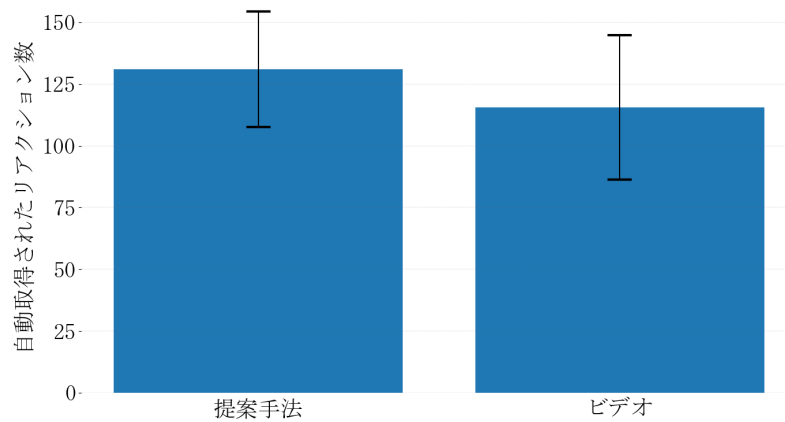


図 5.4: 各条件中に自動検出されたリアクション数.

各条件で自動検出されたリアクションおよび既存手法条件で押されたリアクションの種類と数

提案手法条件で 1 セッション中に自動検出されたリアクションの種類と数の累計データを箱ひげ図にしたものを図 5.5 に示す. 喜び, うなずき, 思考中, 拍手が多く見られ, 驚き, サムズアップ, 挙手はほとんど見られない結果となった.

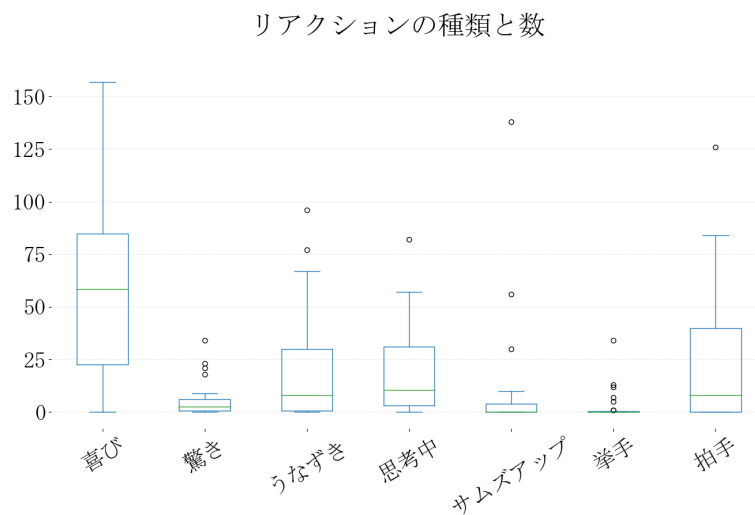


図 5.5: 各条件で自動検出されたリアクションおよび既存手法条件で押されたリアクションの種類と数.

5.3.2 会議に対する評価を測るアンケート

本アンケートは満足度、一体感、リーダーシップ、情報処理の4つの評価指標で会議に対する評価を行う。アンケート結果のデータの平均を棒グラフにしたものを図 5.6 に示す。満足度、一体感、リーダーシップの評価指標のデータにおいて正規性および等分散性が認められたため、二元配置分散分析を行った。情報処理の評価指標のデータにおいては正規性が認められなかったため、ノンパラメトリック検定を行った。

満足度

二元配置分散分析の結果、手法の主効果が有意 ($p=0.044$) であり、議題の主効果および手法と議題の交互作用は有意でなかった。

一体感

二元配置分散分析の結果、手法の主効果が有意 ($p=0.022$) であり、議題の主効果および手法と議題の交互作用は有意でなかった。

リーダーシップ

二元配置分散分析の結果、すべての条件間に有意差は見られなかった。

情報処理

本指標は参加者が会議中の情報をどの程度理解できていると感じるか、また、情報を活用して会議に影響を及ぼせたと感じるかどうかを測る指標である。ウィルコクソンの符号順位検定を行った結果、手法間および議題間では有意差は見られなかった。

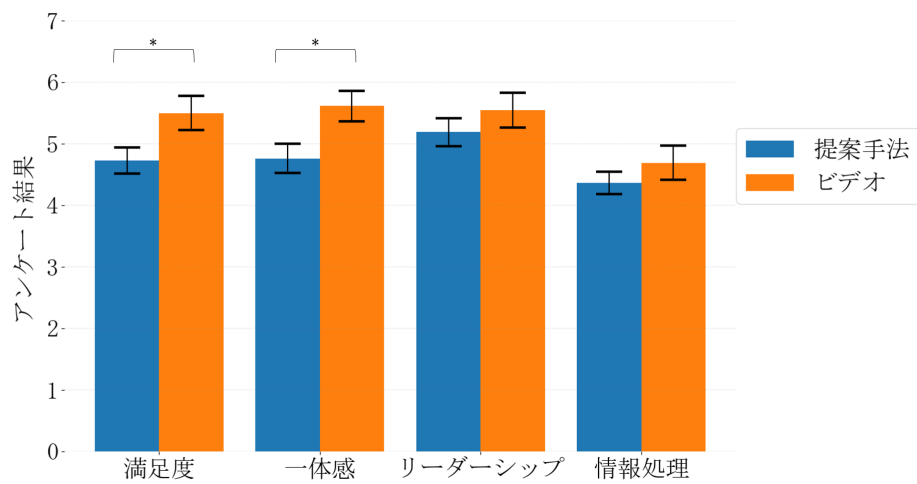


図 5.6: アンケート結果. 有意差が出た条件の間に*を示す.

5.3.3 アイデア数

各セッションにおけるオンラインホワイトボード上に残されたアイデア数の平均データを棒グラフにしたものを図 5.7 に示す. 正規性, 等分散性が認められ, 二元配置分散分析を行った結果, すべての条件間および議題間において有意差は見られなかった.

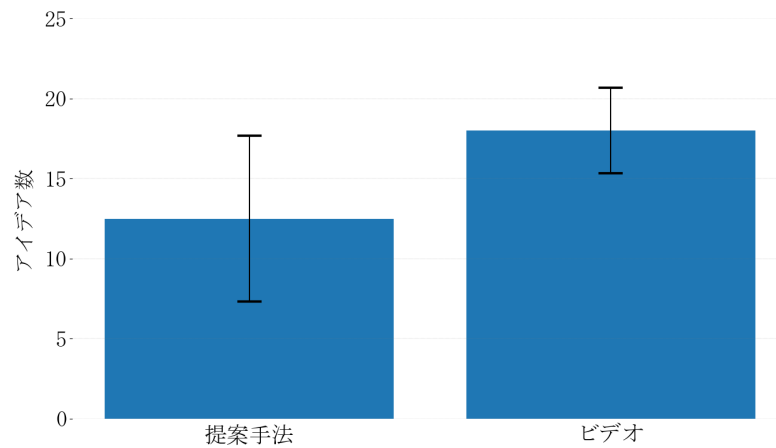


図 5.7: オンラインホワイトボード上に残されたアイデア数.

5.3.4 実験後アンケート

条件の好みについて

“提案手法とビデオのどちらを好むか”という質問に対する回答結果を円グラフにしたものを図 5.8 に示す。提案手法を好んだ参加者からは、“直接相手の顔を見る必要がないため、心理的負担が小さかった”、“一斉に喜びの絵文字が出たときに一体感を感じた”といった意見が見られた。ビデオを好んだ参加者からは、“リアクション以外の動きを伝えることができる”、“ビデオの方が相手の感情を理解できた”、“ビデオは絵文字に比べて情報過多になるかと思ったが、実際はそうでもなかった”等を理由として挙げていた。

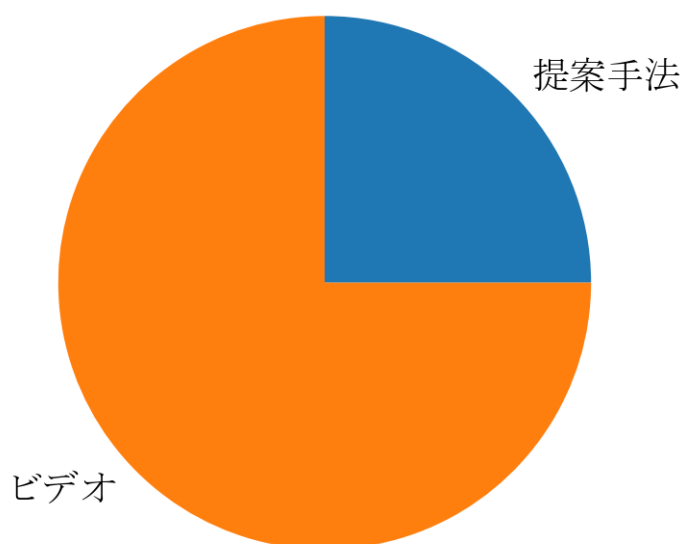


図 5.8: 条件の好みについてのアンケート結果。

入力したいと思ったが、入力できなかったリアクション

回答の結果を図 5.9 に示す。うなずき、思考中、サムズアップ、挙手、拍手の MediaPipe を用いて推定を行ったリアクションについて一定数入力できなかったとの回答を得た。MediaPipe で手ではない部分が手と誤認識される場合がある等、実装上動作を安定させることができなかったことが問題であると考えている。

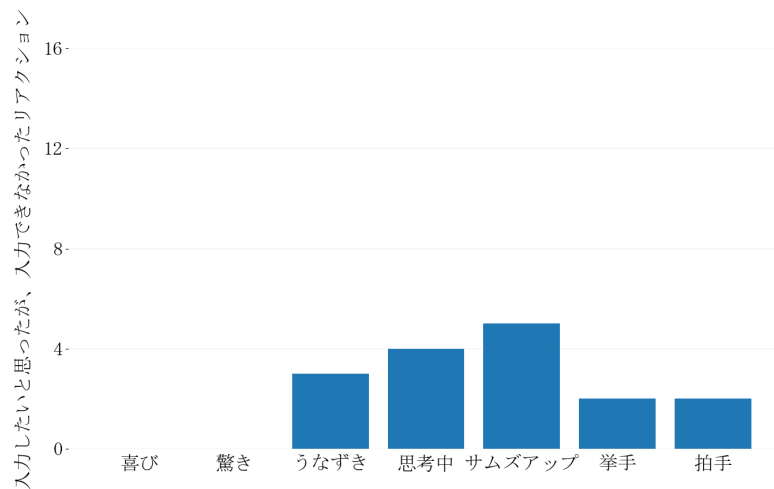


図 5.9: 入力したいと思ったが、入力できなかったリアクション.

入力したいと思ってなかったが、入力されたリアクション

回答の結果を図 5.10 に示す. 喜び, 拍手に関して半数程度の参加者から, 入力したいと思ってなかったが, 入力されたと回答された. 喜びに関しては, 判定の閾値が緩く, 多少の口角の変化程度で判定されていた可能性がある. また, 拍手は 5.3.4 と同様, 手に関する処理を安定させることができなかったからであると考えている.

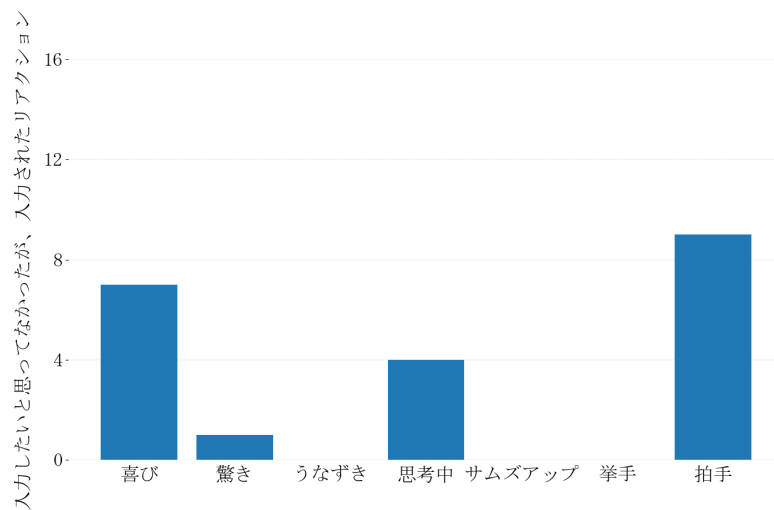


図 5.10: 入力したいと思ってなかったが、入力されたリアクション.

実際に使用したかどうかに関わらず、会議相手に伝えたいと思う機会があったリアクション

回答の結果を図 5.11 に示す。喜び、うなずき、思考中、サムズアップが比較的多くの参加者に必要とされていた。一方で驚き、挙手、拍手は多くの参加者が会議中に用いる機会がなかったと回答した。

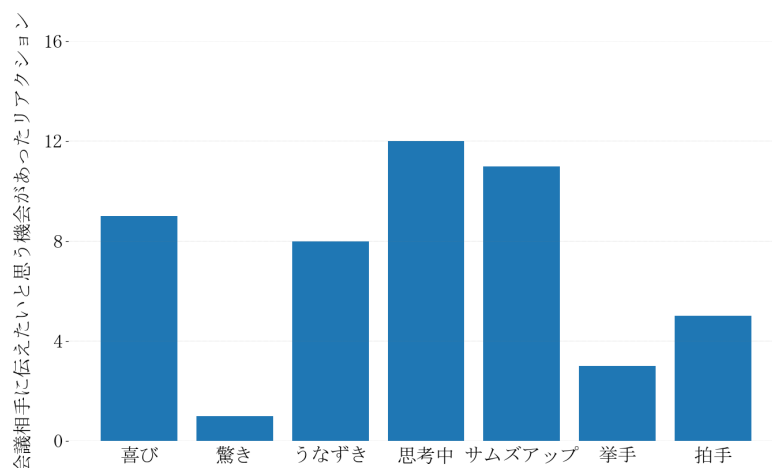


図 5.11: 実際に使用したかどうかに関わらず、会議相手に伝えたいと思う機会があったリアクション。

5.3.5 半構造化インタビュー

絵文字およびビデオのサイズについて

絵文字のサイズに関しては、多くの参加者が現在よりも大きい方がよい、少数の参加者は現在のサイズのままでよいという意見だった。大きい方がよいと答えた参加者からは、“もっと目立たせてよい”、“意識しないと見れない、もっと無意識に知りたい”といった意見が見られた。現在のままでよいと答えた参加者からは、“これ以上大きくしても得られる情報量は変わらない”、“十分目に入った”といった意見が見られた。

ビデオのサイズに関しては、多くの参加者が現在のサイズが適切である、一部の参加者が現在よりも大きい方がよいという意見だった。適切であると答えた参加者からは、“一目で相手の状態を知ることが出来る”、“よく見れば表情もわかるから十分な大きさである”、“これ以上小さいとよくわからない、これ以上大きいと邪魔になる、ちょうどよい大きさだと感じた”といった意見が見られた。大きい方がよいと答えた参加者からは、“相手の表情も見たいから、もう少し大きい方がうれしい”、“邪魔にならないよう注意が必要だが、まだ大きくても大丈夫だと思う”といった意見が見られた。

会議の前後半での違い

絵文字およびビデオの扱いについて前後半で変わったかどうかを質問したところ、絵文字およびビデオをよく見たと答えた方が前半の参加者と後半の参加者で分かれた。前半のアイデアを出す段階の方が絵文字およびビデオを見ている人からは、“自分の意見に対する他の人の反応が気になってよく見ていた”といった意見が見られた。後半の一番良いアイデアを話し合う段階の方が絵文字およびビデオを見ている人からは、“他の人と意見を一致させることが出来ているか確認するため見た”といった意見が見られた。また、その他の意見として、“前半で誰もしゃべらない時間があつたときに思考中が表示されていると安心した”、“後半でサムズアップが扱いやすく良かった”といった意見が見られた。

リアクションの種類

“喜び、うなずき、思考中は自分が出したいと思った場面が多くあつたし、相手のリアクションが見えたらうれしかった”、“驚きは使う場面がなかった”、“挙手は使わなくても発言が出来る環境だったため使わなかった”、“笑いとしての喜び、拍手などの音声から読み取れるリアクションは、音声で十分把握できるため、リアクションとして必要ないと感じた”、“思考中に関しては前回の実装（画面外に顔を向ける）の方が顔だけで完結していて出しやすく良かった”といった意見が見られた。

5.3.6 考察

5.3.1 項、各条件中に自動検出されたリアクション数より、H1 は支持されなかった。仮説ではユーザに自動検出であることを伝えることでユーザが意図的に入力を行おうとし、提案手法条件時に自動検出されるリアクション数が増加すると考えていた。原因として、実験1と同様会議内容に集中するためにリアクションを取ることに意識を割くことができなかったことが考えられる。

5.3.2 項、満足度より、H2 は支持されなかった。仮説ではリアクションを提示することにより、会議中に得られる他の参加者の非言語情報がビデオと同等に近づくことにより、会議の満足度は差が見られなくなると考えていた。原因として、リアクションの種類の選定に問題があつたと考えられる。実験後アンケートおよび半構造化インタビューで得られた意見の中にも、“ビデオの方が相手の感情を理解できた”、“笑いとしての喜び、拍手などの音声から読み取れるリアクションは、音声で十分把握できるため、リアクションとして必要ないと感じた”等が見られた。参加者が他の参加者の映像から得ている情報を解明し、それをリアクションに落とし込むことができれば、H2 が支持される結果となる可能性がある。

5.3.3 項より、アイデア数に有意差は見られなかった。ただし、平均値で見ればビデオ条件の方が6個程度アイデア数が多く、アイデアが同等となるという H3 が支持されたとは言えない。今回の実験では4グループのみだったが、実験参加者を増やし、データの母数を増やすことで有意差が見られる可能性がある。

5.3.1 項、各条件で自動検出されたリアクションおよび既存手法条件で押されたリアクションの種類と数よりリアクションの種類に関して、喜び、うなずき、思考中、拍手が多く見られた。ただし、このうち喜びおよび拍手に関しては5.3.4 項入力したいと思ってなかったが、入力されたリアクションより、意図せず入力されてしまった可能性がある。また、5.3.4 項実際に使用したかどうかに関わらず、会議相手に伝えたいと思う機会があったリアクションより、上記以外にサムズアップが会議中に伝えたい機会があったリアクションとして挙げられていた。しかし、回数として多く見られたわけではないため、機会はあったものの回数が少なかったか、行動として移すまでではなかったかのいずれかであると考えられる。

5.3.2 項、一体感より、一体感は提案手法-ビデオ間で有意差が見られた。参加者はビデオにより他の参加者の顔が直接見えることで、一体感を感じていると考えている。実験後アンケートおよび半構造化インタビューでも、“ビデオで他の参加者の顔が見れると安心する”といった意見がみられた。

5.3.2 項、リーダーシップより、リーダーシップにおいて有意差は見られなかった。本実験で行われた会議は、すべての参加者が対等な立場で意見を出し合うものであり、特定の個人によって主導で行われるものではない。また、実験前に注意事項として、発言量、作業量が偏らないよう、議長、書記等の明確な役割を持つことを禁止していた。そのため、リーダーシップの評価で有意差が見られなかったと考えられる。

5.3.2 項、情報処理より、情報処理は有意差は見られなかった。本指標は参加者が会議中の情報をどの程度理解できていると感じるか、また、情報を活用して会議に影響を及ぼせたと感じるかどうかを測る指標である。前述の通り満足度、一体感についてはビデオ条件の方が高かったが、そのような会議に対する印象の改善は、会議でやり取りされるアイデア、意見等の情報の量や質に影響しない程度であり、結果として情報処理の評価で有意差が見られなかったと考える。

5.4 議論

5.4.1 本実験の制約、今後の課題

自動検出の認識精度

半構造化インタビューおよび実験後アンケートの中で、一部の参加者から意図的に出すことができないリアクションおよび意図していないリアクションの出力が起きたという意見が見られた。また、参加者の中には認識精度に対する不安を感じるという意見も見られた。特にうなずき、思考中、サムズアップ、挙手、拍手の MediaPipe を用いたリアクションを取ることができない、喜びおよび拍手のリアクションが意図せず出力されていたことがアンケート結果から明らかになった。ただし、ユーザごとの個人差および背景等の環境の差により認識精度に影響が生じている可能性がある。よって今後は face-api.js [jus], MediaPipe [Good] 以外の識別手法を用いる、より適切な閾値処理を行う等精度向上のための更新を行う。

実験に用いたモニタの画面サイズ

本研究の実験では、実験参加者間で環境の差が生じないようにするため、23.8inch モニタを用いて実験を行った。しかし、実際にオンラインホワイトボードを用いたオンライン会議を行う際に、外部モニタを用いずラップトップコンピュータのみで参加する場合も考えられる。その際に、実験よりも小さな画面サイズでオンラインホワイトボードの作業を求められるため、ビデオと比較し、より小さな表示領域で済ませることができる提案手法に有利に働く可能性があると考えられる。今後、絵文字およびビデオの表示サイズとともに調査が必要である。

実験参加者間の関係

本研究の実験では、実験を円滑に進めるため、あらかじめ知り合いである4人組で会議を行ってもらった。また、その4人は学生同士であり、学年の差はあれど立場に大きな差はなかったと考えられる。しかし、会議を行う人同士が初対面であったり、大きな立場の違いがあった場合、心理的負担の観点から、提案手法の顔映像を映さなくてよいという点が有利に働く可能性がある。今後、参加者間の関係が提案手法の効果に影響を与えるかの調査が必要である。

5.4.2 今後の展望

リアクションの種類の設定

本実験ではリアクションは喜び、驚き、うなずき、思考中、サムズアップ、挙手、拍手の7種類を選定した。実験1の結果および参加者からの意見を参考に選定を行ったが、結果として驚き、挙手、拍手に関しては会議中あまり用いられないことが分かった。ただし、この結果は会議の内容や人数に左右される可能性がある。例えば、会議人数が多くなることにより、発言に対するハードルが上がり挙手が必要とされる、決定会議等の物事を決める会議で最終決定がされた際に拍手が必要とされる等が考えられる。今後、会議の内容、人数による需要の差も含めた、必要とされるリアクションの種類についての調査が必要である。

絵文字およびビデオの表示サイズ

本実験では、絵文字は50×50px、ビデオは100×100pxで表示をした。半構造化インタビューでは、絵文字に関しては多くの参加者が現在よりも大きい方がよい、少数の参加者は現在のサイズのままでよいと回答した。ビデオに関しては、多くの参加者が現在のサイズが適切である、一部の参加者が現在よりも大きい方がよいと回答した。また、本研究の実験では23.8inchのモニタを用いたが、さらに画面サイズが小さく、または大きくなった際に適切なサイズはどのように変化するか不明である。絵文字およびビデオに関して、ホワイトボードの作業の邪魔にならず、他の参加者の非言語情報を読み取れる適切なサイズを調査する必要がある。

第6章 おわりに

本研究では、オンラインホワイトボードを用いるオンライン会議において、ビデオ共有を用いず、顔映像からリアクション推定を行い、絵文字化してカーソル位置に提示する手法を提案した。リアクションは表情、頭部姿勢、手の姿勢および位置を推定に用いて、実験1では喜び、悲しみ、驚き、恐れ、怒り、嫌悪、うなずき、思考中を、実験2では喜び、驚き、うなずき、思考中、サムズアップ、挙手、拍手を提示した。本研究では、提案手法の効果を検証するため実験を2回行った。実験は両方とも4人1グループで行った。1回目の実験では、一般的に広く用いられている、リアクションボタンを押すことによる共有を行う既存手法および画面端にビデオを表示する手法の2つの手法と比較した。実験の結果、既存手法と比較して提案手法の方が多くのリアクションが共有され、会議に対する満足度の向上が見られた。また、提案手法とビデオ間には満足度の差は見られず、アイデア数ではすべての条件間で差は見られなかった。2回目の実験では、従来のビデオの表示方法よりも優れた表示方法と考えられる、カーソル位置にビデオを表示する手法と比較した。実験の結果、会議に対する満足度は提案手法よりもビデオの方が高くなった。また、アイデア数では差は見られなかった。今後の展望として、ビデオの代替となりうるリアクションの選定や、ディスプレイサイズ等の様々な環境下での適切な絵文字の表示サイズについて検討することが挙げられる。

謝辞

本研究を進めるにあたり、志築文太郎先生、高橋伸先生、川口一画先生には多大なご意見とご指導を頂きました。心から感謝いたします。特に、実質指導教員である川口一画先生には、研究の進め方、論文執筆、発表の方法といった研究に関することのみならず、日常生活における研究者の心構えなど多岐に渡るご指導を頂きました。先生の熱心なご指導のおかげで、これまでの研究成果の発表および本論文の執筆ができました。ここに深く感謝の意を表します。

インタラクティブプログラミング研究室の後輩、同輩、先輩方には研究生活においてお世話になりました。特に COMMUNICATION チームの皆様には、チームゼミにおけるご意見や論文の添削といった研究に関する多くのご支援に加えて、研究室においても研究の進め方や研究生活について多くのご助言を頂きました。研究室の皆様のおかげで、3年間充実した研究生活を送ることができました。深く感謝いたします。

最後に、学生生活においてお世話になった皆様、そして、私の学生生活を支えて頂いた家族に深く感謝を申し上げます。

参考文献

- [BYMS06] Jeremy N Bailenson, Nick Yee, Dan Merget, and Ralph Schroeder. The Effect of Behavioral Realism and Form Realism of Real-Time Avatar Faces on Verbal Disclosure, Nonverbal Disclosure, Emotion Recognition, and Copresence in Dyadic Interaction. *Presence: Teleoperators and Virtual Environments*, Vol. 15, No. 4, pp. 359–372, 08 2006.
- [BZLM18] Tadas Baltrusaitis, Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency. Openface 2.0: Facial behavior analysis toolkit. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pp. 59–66, 2018.
- [Cor] NTT Communications Corporation. Skyway. <https://skyway.ntt.com/>.
- [DJMW98] OWEN DALY-JONES, ANDREW MONK, and LEON WATTS. Some advantages of video conferencing over high-quality audio conferencing. *Int. J. Hum.-Comput. Stud.*, Vol. 49, No. 1, p. 21–58, jul 1998.
- [EF78] P Ekman and WV Friesen. Facial coding action system (facs): A technique for the measurement of facial actions, 1978.
- [EFE82] P. Ekman, W.V. Friesen, and P. Ellsworth. What are the relative contributions of facial behavior and contextual information to the judgment of emotion? In *Emotion in the human face*, pp. 111–127, 1982.
- [Fig] Figma. Figma. <https://www.figma.com/>.
- [Gooa] Google. Firebase realtime database. <https://firebase.google.com/docs/database>.
- [Goob] Google. Google jamboard. <https://edu.google.com/jamboard/>.
- [Gooc] Google. Google meet. <https://workspace.google.com/intl/meet/products/meet/>.
- [Good] Google. Mediapipe. <https://developers.google.com/mediapipe>.
- [GT31] F.L. Goodenough and M.A. Tinker. The relative potency of facial expression and verbal description of stimulus in the judgment of emotion. *Journal of Comparative Psychology*, Vol. 12, No. 4, pp. 365–370, 1931.

- [jus] justadudewhohacks. face-api.js. <https://github.com/justadudewhohacks/face-api.js>.
- [KAL⁺17] Tero Karras, Timo Aila, Samuli Laine, Antti Herva, and Jaakko Lehtinen. Audio-driven facial animation by joint end-to-end learning of pose and emotion. Vol. 36, No. 4, 2017.
- [Mic] Microsoft. Microsoft Teams <https://www.microsoft.com/microsoft-teams/group-chat-software>.
- [mir] miro. miro. <https://miro.com/>.
- [MP59] A. Meadow and S. J Parnes. Evaluation of training in creative problem solving. *Journal of Applied Psychology*, Vol. 43, No. 3, pp. 189–194, 1959.
- [MUR] MURAL. Mural. <https://www.mural.co/>.
- [NSI⁺21] Kosaku Namikawa, Ippei Suzuki, Ryo Iijima, Sayan Sarcar, and Yoichi Ochiai. Emojicam: Emoji-assisted video communication system leveraging facial expressions. In Masaaki Kurosu, editor, *Human-Computer Interaction. Design and User Experience Case Studies*, pp. 611–625. Springer International Publishing, 2021.
- [OOM95] Judith S. Olson, Gary M. Olson, and David K. Meader. What mix of video and audio is useful for small groups doing remote real-time design work? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '95, p. 362–368. ACM Press/Addison-Wesley Publishing Co., 1995.
- [OSM⁺20] Kotaro Oomori, Akihisa Shitara, Tatsuya Minagawa, Sayan Sarcar, and Yoichi Ochiai. A preliminary study on understanding voice-only online meetings using emoji-based captioning for deaf or hard of hearing users. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '20. Association for Computing Machinery, 2020.
- [Pit] Pitch. Pitch. <https://pitch.com/>.
- [R C16] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, 2016.
- [Spa] SpatialChat. Spatialchat. <https://www.spatial.chat/>.
- [W.M.08] W.M. Post, M.A.A. Huis in 't Veld, van den Boogaard. Evaluating meeting support tools. In *Personal and Ubiquitous Computing*, 第 12 卷, pp. 223—235, 2008.
- [ZM89] F Zijlstra and Th Meijman. Het meten van mentale inspanning met behulp van een subjectieve methode (measurement of mental effort with a subjective method). *Mentale belasting en werkstress. Een arbeidspsychologische benadering*, pp. 42–61, 1989.

- [Zoo] Zoom. Zoom Meetings
- [会議 08] 会議の進め方. 日経文庫. 日本経済新聞出版社, 2008.
- [岸本 22] 岸本諒大, 土田修平, 寺田努, 塚本昌彦. Faceshow: 顔の部分表示による遠隔会議システム. 日本ソフトウェア科学会第 30 回インタラクティブシステムとソフトウェアに関するワークショップ (WISS2022) 論文集, pp. 31–37, dec 2022.
- [佐藤 22] 佐藤花菜, 石川知一. オンラインコミュニケーションにおける非言語情報可視化の実装と影響の調査. 第 84 回全国大会講演論文集, 第 2022 巻, pp. 157–158, feb 2022.
- [小俣 09] 小俣昌樹, 深澤寛和, 今宮淳美. Rj-005 仮想影カーソルによる遠隔協調作業時のコミュニケーション支援 (ヒューマンコミュニケーション&インタラクション, 査読付き論文). 情報科学技術フォーラム講演論文集, 第 8 巻, pp. 57–60, aug 2009.
- [長尾 21] 長尾優花, 北野圭介, 大島登志一, 望月茂徳. 3d アバタを利用した遠隔授業での頭部動作の認識を用いた意思表示インタフェース. エンタテインメントコンピューティングシンポジウム論文集, 第 2021 巻, pp. 346–350, aug 2021.

著者論文リスト

本論文に関する論文および発表

- 論文誌原著論文

1. 大山 智弘, 川口 一画. オンラインホワイトボードを用いた創造会議における自動検出による非言語情報伝達. ヒューマンインタフェース学会論文誌, (Vol.26, No.1 採録予定)

その他論文

- 査読なし国内会議論文

1. 大山 智弘, 川口 一画. アドベンチャーゲームにおけるユーザ体験向上のための感情推定技術を用いた入力手法. 第 199 回 HCI 研究会, 情報処理学会, 2022 年 8 月, 8pages.
2. 飯塚 陸斗, 大山 智弘, 杉山 将利, 下田 康太, 田之頭 吾音, 井原 圭一, 石橋 遼, 川口 一画. 動的音高変化による方向提示手法設計のための基礎的調査. インタラクション 2023, 情報処理学会, 2023 年 3 月, 5pages.

付 録A 実験におけるアンケート

4, 5 章にて述べた実験に用いたアンケートを示す.

A.1 タスク後アンケート

実験 1 で用いたものを記載する. 実験 2 では実験条件の項目のみ変更して使用した.

タスク後アンケート

1:同意しない⇔7:同意する

* 必須の質問です

1。 参加者ID *

2。 年齢 *

3。 実験条件 *

1つだけマークしてください。

☐ リアクションボタン

☐ 自動取得

☐ ビデオ

4。 この会議に満足した *

1つだけマークしてください。

1 2 3 4 5 6 7

☐ ☐ ☐ ☐ ☐ ☐ ☐

5。 会議相手を協力的に感じた *

1つだけマークしてください。

1 2 3 4 5 6 7

☐ ☐ ☐ ☐ ☐ ☐ ☐

6。 会議が良い方向に進んだ *

1つだけマークしてください。

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

7。 すべての情報が使われていた *

1つだけマークしてください。

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

8。 議題に対してよいアイデアを出せると信じていた *

1つだけマークしてください。

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

9。 全体を通して満足した *

1つだけマークしてください。

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

10。 会議相手を信頼できると感じた *

1つだけマークしてください。

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

11. 会議におけるアイデア出しの進行に満足している *

1つだけマークしてください。

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

12. 自分の貢献が会議の意思決定に影響したと考えている *

1つだけマークしてください。

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

13. 会議相手と一緒にいると楽しい *

1つだけマークしてください。

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

14. すべての会議参加者に貢献をする機会が十分にあった *

1つだけマークしてください。

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

15. 会議相手にイラつきを覚えた *

1つだけマークしてください。

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

16. 会議相手に優しさを感じた *

1つだけマークしてください。

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

17. 情報が多すぎた *

1つだけマークしてください。

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

18. 会議参加者全員が十分な配慮を受けた *

1つだけマークしてください。

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

このコンテンツは Google が作成または承認したものではありません。

Google フォーム

A.2 追加アンケート

実験 1 後に行った追加調査で用いたアンケートを記載する.

追加アンケート

よくわからない箇所があったら大山までお願いします

1. 入力したいと思ったが、入力できなかったリアクション

当てはまるものをすべて選択してください。

- ☐ 喜び
- ☐ 悲しみ
- ☐ 驚き
- ☐ 恐れ
- ☐ 怒り
- ☐ 嫌悪
- ☐ うなずき
- ☐ 思考中

2. 入力したいと思ってなかったが、入力されたリアクション

当てはまるものをすべて選択してください。

- ☐ 喜び
- ☐ 悲しみ
- ☐ 驚き
- ☐ 恐れ
- ☐ 怒り
- ☐ 嫌悪
- ☐ うなずき
- ☐ 思考中

3. 実際に使用したかどうかに関わらず、会議相手に伝えたいと思う機会があったリアクション

当てはまるものをすべて選択してください。

- ☐ 喜び
- ☐ 悲しみ
- ☐ 驚き
- ☐ 恐れ
- ☐ 怒り
- ☐ 嫌悪
- ☐ うなずき
- ☐ 思考中

A.3 実験後アンケート

実験 2 で用いた実験後アンケートを記載する.

追加アンケート

よくわからない箇所があったら大山までお願いします

* 必須の質問です

1. どちらの手法が好みでしたか？ *

1つだけマークしてください。

☐ 絵文字

☐ ビデオ

2. 絵文字手法のとき、入力したいと思ったが、入力できなかったリアクション

当てはまるものをすべて選択してください。

☐ 喜び

☐ 驚き

☐ うなずき

☐ 思考中

☐ サムズアップ

☐ 挙手

☐ 拍手

3. 絵文字手法のとき、入力したいと思ってなかったが、入力されたリアクション

当てはまるものをすべて選択してください。

☐ 喜び

☐ 驚き

☐ うなずき

☐ 思考中

☐ サムズアップ

☐ 挙手

☐ 拍手

4. 手法、実際に使用したかどうかに関わらず、会議相手に伝えたいと思う機会があったリ
アクション

当てはまるものをすべて選択してください。

- ☐ 喜び
- ☐ 驚き
- ☐ うなずき
- ☐ 思考中
- ☐ サムズアップ
- ☐ 挙手
- ☐ 拍手

5. 自由意見

このコンテンツは Google が作成または承認したものではありません。

Google フォーム